

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number  
**WO 03/014340 A2**

(51) International Patent Classification<sup>7</sup>: C12N 9/00      (74) Agent: GROS, Florent; Novartis AG, Corporate Intellectual Property, Patent & Trademark Department, CH-4002 Basel (CH).

(21) International Application Number: PCT/EP02/08654

(22) International Filing Date: 2 August 2002 (02.08.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/309,957                    3 August 2001 (03.08.2001) US

(71) Applicant (for all designated States except AT, US): NOVARTIS AG [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH).

(71) Applicant (for AT only): NOVARTIS PHARMA GMBH [AT/AT]; Brunner Strasse 59, A-1230 Vienna (AT).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ATADJA, Peter, Wisdom [CA/US]; 18 Eastbrook Road, Parsippany, NJ 07054 (US). CUETO, Maria [US/US]; 99 Clifton Terrace, Weehawken, NJ 07086 (US). GAO, Lin [US/US]; 8 Millstone Drive, Morris Plains, NJ 07950 (US).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LT, LU, LV, MA, MD, MK, MN, MX, NO, NZ, OM, PH, PL, PT, RO, RU, SE, SG, SI, SK, TJ, TM, TN, TR, TT, UA, US, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.



**WO 03/014340 A2**

(54) Title: HISTONE DEACETYLASE-RELATED GENE AND PROTEIN

(57) Abstract: Disclosed is an HDAC related gene and gene product. In particular, the invention relates to a protein that is highly homologous to known HDACs and referred to herein as HDAC10, nucleic acid molecules that encode such a protein, antibodies that recognize the protein, and methods for diagnosing conditions related to abnormal HDAC10 activity or gene expression.

### HISTONE DEACETYLASE - RELATED GENE AND PROTEIN

This invention relates to a histone deacetylase gene and gene product. In particular, the invention relates to a protein that is highly homologous to known mammalian histone deacetylases (HDACs), nucleic acid molecules that encode such a protein, antibodies that recognize the protein, and methods of use which include assays screening for modulators of HDAC activity and for diagnosing conditions related to abnormal HDAC activity, including, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response or psoriasis.

#### **BACKGROUND**

Histone acetylation is a major regulatory mechanism that modulates gene expression by altering the accessibility of transcription factors to DNA. Acetylation of histones is a reversible modification of the free ε-amino group of lysine that occurs during the assembly of nucleosomes and during DNA synthesis.

HDACs have been shown to play an important role in the regulation of transcription. HDACs function as components of complexes that are involved in transcriptional repression. This is mediated through interactions of HDACs with multi-protein complexes and requires deacetylase activity. Changes in histone acetylation levels also occur during transcriptional activation and silencing. Acetylation of histones is generally associated with transcriptional activity, whereas deacetylation is associated with transcriptional repression.

HDAC complexes may contain the co-repressor mSin3A and mSin3A-associated proteins, silencing mediators NcoR and SMRT, transcriptional repressors, Rb-like proteins p107 and p130, Rb-associated proteins, nuclear hormone receptors, nucleosome remodeling factors, methyl-binding proteins, DNA repair machinery proteins, and the like. Furthermore, HDAC1 has been found to bind directly to YY1 and Sp1 and HDACs 4 and 5 bind to MEF2. In addition, HDACs have been found together in complexes.

Two distinct classes of yeast histone deacetylases have been identified based upon size and sequence. Yeast class I HDACs include Rpd3, Hos1p, and Hos2p. Class II contains yeast HDA1p.

Furthermore, members of these two classes were found to form different complexes. Human HDACs have been classified based upon their similarity to yeast sequences. Class I human HDACs include HDACs1-3 and 8. Class II HDACs include HDACs 4-7. The deacetylase core of class I HDACs reside in the first ~390 amino acids. Class II HDAC catalytic domains are located in the C-terminal of these peptides, with the exception of HDAC6 that contains a second catalytic domain in the N-terminus. Here we report the isolation and characterization of a new HDAC, referred to herein as HDAC10.

An important approach that has been used to study the function of chromatin acetylation is the use of specific inhibitors of histone deacetylase. Several classes of compounds have been identified that inhibit HDAC. Histone deacetylase inhibitors have been found to have anti-proliferative effects, including induction of G1/S and G2/M cell cycle arrest, differentiation and apoptosis of transformed and normal cells and reversal of transformation. These effects, along with the presence of HDAC in complexes with fusions of unliganded retinoic acid receptors PML-RAR $\alpha$  and PLZF-RAR $\alpha$  indicate a role for HDACs in tumorigenicity. Furthermore, histone deacetylase inhibitors, phenylbutyrate and trichostatin A have shown promise in the treatment of promyelocytic leukemia and several other HDAC inhibitors are being studied as treatments for cancers.

#### SUMMARY OF THE INVENTION

The present invention relates to a novel histone deacetylase designated HDAC10.

In a first aspect, the invention provides an isolated polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1. Furthermore, the invention provides an isolated polypeptide consisting of an amino acid sequence as set forth in SEQ ID NO:1. The amino acid sequence as set forth in SEQ ID NO:1 shows a considerable degree of homology to that of known members of the family of HDACs in the catalytic domain. For convenience, the polypeptide consisting of the amino acid sequence as set forth in SEQ ID NO:1 will be designated as histone deacetylase 10 or HDAC10. Fragments of the isolated polypeptide having an amino acid sequence as set forth in SEQ ID NO:1 also form a part of the present invention. Preferably, fragments will encompass the catalytic domain, which is predicted to exist between amino acid number 15 to 323. In accordance with this aspect of the invention there are provided novel polypeptides of human origin as well as biologically, diagnostically or therapeutically useful fragments, variants and derivatives thereof, variants and derivatives of the fragments, and analogs of the foregoing.

In a second aspect, the invention provides an isolated DNA comprising a nucleotide sequence that encodes a polypeptide as mentioned above. In particular, the invention provides (1) an isolated DNA comprising the nucleotide sequence as set forth in SEQ ID NO:2; (2) an isolated DNA comprising the nucleotide sequence set forth in SEQ ID NO:3; (3) an isolated DNA capable of hybridizing under high stringency conditions to the nucleotide sequence set forth in SEQ ID NO:2; and (4) an isolated DNA comprising the nucleotide sequence set forth in SEQ ID NO:4. Also provided are nucleic acid sequences comprising at least about 15 bases, preferably at least about 20 bases, more preferably a nucleic acid sequence comprising about 30 contiguous bases of SEQ ID NO:2 or SEQ ID NO:3. Also within the scope of the present invention are nucleic acids that are substantially similar to the nucleic acid with the nucleotide sequence as set forth in SEQ ID NO:2 or SEQ ID NO:3. In a preferred embodiment, the isolated DNA takes the form of a vector molecule comprising at least a fragment of a DNA of the present invention, in particular comprising the DNA consisting of a nucleotide sequence as set forth in SEQ ID NO:2 or SEQ ID NO:3.

A third aspect of the present invention encompasses a method for the diagnosis of conditions associated with abnormal regulation of gene expression which includes, but is not limited to, conditions associated with abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, or psoriasis in a human which comprises detecting abnormal transcription of messenger RNA transcribed from the natural endogenous human gene encoding the novel polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1 in an appropriate tissue or cell from a human, wherein such abnormal transcription is diagnostic of the human's affliction with such a condition. In particular, the said natural endogenous human gene encoding the novel polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1 comprises the genomic nucleotide sequence set forth in SEQ ID NO:4. In one embodiment of the present invention, the diagnostic method comprises contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from that tissue or cell with an isolated nucleotide sequence of at least about 15 - 20 nucleotides in length that hybridizes under high stringency conditions with the isolated nucleotide sequence encoding the novel polypeptide having an amino acid sequence set forth in SEQ ID NO:1.

Another embodiment of the assay aspect of the invention provides a method for the diagnosis of a condition associated with abnormal HDAC10 activity in a human, which comprises measuring the level of deacetylase activity in a certain tissue or cell from a human suffering from such a condition, wherein the presence of an abnormal level of deacetylase activity, relative to the level

thereof in the respective tissue or cell of a human not suffering from a condition associated with abnormal HDAC activity, is diagnostic of the human's suffering from said condition.

In accordance with one embodiment of this aspect of the invention there are provided anti-sense polynucleotides that can regulate transcription of the gene encoding the novel HDAC10; in another embodiment, double stranded RNA is provided that can regulate the transcription of the gene encoding the novel HDAC10.

Another aspect of the invention provides a process for producing the aforementioned polypeptides, polypeptide fragments, variants and derivatives, fragments of the variants and derivatives, and analogs of the foregoing. In a preferred embodiment of this aspect of the invention there are provided methods for producing the aforementioned HDAC10 comprising culturing host cells having incorporated therein an expression vector containing an exogenously-derived nucleotide sequence encoding such a polynucleotide under conditions sufficient for expression of the polypeptide in the host cell, thereby causing expression of the polypeptide, and optionally recovering the expressed polypeptide. In a preferred embodiment of this aspect of the present invention, there is provided a method for producing polypeptides comprising or consisting of an amino acid sequence as set forth in SEQ ID NO:1, which comprises culturing a host cell having incorporated therein an expression vector containing an exogenously-derived polynucleotide encoding a polypeptide comprising or consisting of an amino acid sequence as set forth in SEQ ID NO:1, under conditions sufficient for expression of such a polypeptide in the host cell, thereby causing the production of an expressed polypeptide, and optionally recovering the expressed polypeptide. Preferably, in any of such methods the exogenously derived polynucleotide comprises or consists of the nucleotide sequence set forth in SEQ ID NO:2, the nucleotide sequence set forth in SEQ ID NO:3, or the nucleotide sequence set forth in SEQ ID NO:4. In accordance with another aspect of the invention there are provided products, compositions, processes and methods that utilize the aforementioned polypeptides and polynucleotides for, *inter alia*, research, biological, clinical and therapeutic purposes.

In certain additional preferred embodiments of this aspect of the invention there is provided an antibody or a fragment thereof which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1, i.e., HDAC10. In certain particularly preferred embodiments in this regard, the antibodies are highly selective for human HDAC10 polypeptides or portions of human HDAC10 polypeptides.

In a further aspect, an antibody or fragment thereof is provided that binds to a fragment or portion of the amino acid sequence set forth in SEQ ID NO:1.

In another aspect, methods of treating a condition in a subject, wherein the condition is associated with abnormal HDAC10 gene expression, an increase or decrease in the presence of HDAC10 polypeptide in a subject, or an increase or decrease in the activity of HDAC10 polypeptide, by the administration of an effective amount of an antibody that binds to a polypeptide with the amino acid sequence set out in SEQ ID NO:1, or a fragment or portion thereof to the subject are provided. Also provided are methods for the diagnosis of a disease or condition associated with abnormal HDAC10 gene expression or an increase or decrease in the presence of the HDAC10 in a subject, or an increase or decrease in the activity of HDAC10 polypeptide.

In yet another aspect, the invention provides host cells which can be propagated *in vitro*, preferably vertebrate cells, in particular mammalian cells, or bacterial cells, which are capable upon growth in culture of producing a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1 or fragments thereof, where the cells contain transcriptional control DNA sequences, where the transcriptional control sequences control transcription of RNA encoding a polypeptide with the amino acid sequence according to SEQ ID NO:1 or fragments thereof. This includes, but is not limited to, the propagation of HDAC10 in a plasmid and the production of DNA, RNA or protein in human or insect cells or bacteria using the endogenous HDAC10 promoter or any other transcriptional control sequence.

In yet another aspect of the present invention there are provided assay methods and kits comprising the components necessary to detect above-normal expression of polynucleotides encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1, or polypeptides comprising an amino acid sequence set forth in SEQ ID NO:1, or fragments thereof, in body tissue samples derived from a patient, such kits comprising e.g., antibodies that bind to a polypeptide comprising an amino acid sequence set forth in SEQ ID NO:1, or to fragments thereof, or oligonucleotide probes that hybridize with polynucleotides of the invention. In a preferred embodiment, such kits also comprise instructions detailing the procedures by which the kit components are to be used.

In another aspect, the invention is directed to use of a polypeptide comprising an amino acid sequence set forth in SEQ ID NO:1 or fragment thereof, polynucleotide encoding such a polypeptide or a fragment thereof, or antibody that binds to said polypeptide comprising an amino acid sequence set forth in SEQ ID NO:1 or a fragment thereof in the manufacture of a medicament to treat diseases associated with abnormal HDAC activity or gene expression.

Another aspect is directed to pharmaceutical compositions comprising a polypeptide comprising or consisting of an amino acid sequence set forth in SEQ ID NO:1 or fragment thereof, a polynucleotide encoding such a polypeptide or a fragment thereof, or antibody that binds to such a polypeptide or a fragment thereof, in conjunction with a suitable pharmaceutical carrier, excipient or diluent, for the treatment of diseases associated with abnormal HDAC activity or gene expression.

In another aspect, the invention is directed to methods for the identification of molecules that can bind to a polypeptide comprising an amino acid sequence set forth in SEQ ID NO:1 and/or modulate the activity of a polypeptide comprising an amino acid sequence set forth in SEQ ID NO:1 or molecules that can bind to nucleic acid sequences that modulate the transcription or translation of a polynucleotide encoding a polypeptide comprising an amino acid sequence set forth in SEQ ID NO:1. Molecules identified by such methods also fall within the scope of the present invention.

In a related aspect, the invention is directed to use of the novel HDAC10 to identify associated proteins in HDAC biologically relevant complexes. At present, the proteins that associate with HDAC10 are not known. However, these may be characterized by determining whether HDAC10 associates with proteins that have been previously shown to interact with other HDACs (see Introduction). For example, components of HDAC10 complexes may be determined using conventional methods, including co-immunoprecipitation.

In yet another aspect, the invention is directed to methods for the introduction of nucleic acids of the invention into one or more tissues of a subject in need of treatment with the result that one or more proteins encoded by the nucleic acids are expressed and or secreted by cells within the tissue.

Other objects, features, advantages and aspects of the present invention will become apparent to those of skill from the following description. It should be understood, however, that the following description and the specific examples, while indicating preferred embodiments of the invention, are given by way of illustration only. Various changes and modifications within the spirit and scope of

the disclosed invention will become readily apparent to those skilled in the art from reading the following description and from reading the other parts of the present disclosure.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows amino acid sequence (SEQ ID NO:1) of HDAC10.

Figure 2 shows the full-length cDNA sequence (SEQ ID NO:2) of HDAC10. The full-length cDNA sequence starts at nucleotide position 1 and ends at nucleotide position 1755.

Figure 3 shows the open reading frame of HDAC10 cDNA sequence (SEQ ID NO:3). The sequence starts at nucleotide position 25 and ends at nucleotide position 1065 as indicated in SEQ ID NO:2.

Figure 4 shows HDAC10 genomic DNA sequence (SEQ ID NO:4).

#### DETAILED DESCRIPTION OF THE INVENTION

In practicing the present invention, many conventional techniques in molecular biology, microbiology, and recombinant DNA are used. These techniques are well known to one of ordinary skill in the art. The following abbreviations used throughout the disclosure are listed herein below: histone deacetylase (HDAC), histone deacetylase-like protein (HDLP)

In its broadest sense, the term "substantially similar", when used herein with respect to a nucleotide sequence, means a nucleotide sequence corresponding to a reference nucleotide sequence, wherein the corresponding sequence encodes a polypeptide having substantially the same structure and function as the polypeptide encoded by the reference nucleotide sequence, e.g. where only changes in amino acids not affecting the polypeptide function occur. Desirably the substantially similar nucleotide sequence encodes the polypeptide encoded by the reference nucleotide sequence. The percentage of identity between the substantially similar nucleotide sequence and the reference nucleotide sequence desirably is at least 80%, more desirably at least 85%, preferably at least 90%, more preferably at least 95%, still more preferably at least 98 or 99%. Sequence comparisons are

carried out using Clustalw (see, for example, Higgins, D.G. et al. *Methods Enzymol.* 266:383-402 (1996)). Clustalw alignments were performed using default parameters.

A nucleotide sequence "substantially similar" to reference nucleotide sequence hybridizes to the reference nucleotide sequence in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO<sub>4</sub>, 1 mM EDTA at 50°C with washing in 2X SSC, 0.1% SDS at 50°C, more desirably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO<sub>4</sub>, 1 mM EDTA at 50°C with washing in 1X SSC, 0.1% SDS at 50°C, more desirably still in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO<sub>4</sub>, 1 mM EDTA at 50°C with washing in 0.5X SSC, 0.1% SDS at 50°C, preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO<sub>4</sub>, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 50°C, more preferably in 7% sodium dodecyl sulfate (SDS), 0.5 M NaPO<sub>4</sub>, 1 mM EDTA at 50°C with washing in 0.1X SSC, 0.1% SDS at 65°C, yet still encodes a functionally equivalent gene product.

"Elevated transcription of mRNA" refers to a greater amount of messenger RNA transcribed from the natural endogenous human gene encoding the novel polypeptide of the present invention present in an appropriate tissue or cell of an individual suffering from a condition associated with abnormal HDAC10 activity than in a subject not suffering from such a disease or condition; in particular at least about twice, preferably at least about five times, more preferably at least about ten times, most preferably at least about 100 times the amount of mRNA found in corresponding tissues in humans who do not suffer from such a condition. Such elevated level of mRNA may eventually lead to increased levels of protein translated from such mRNA in an individual suffering from a condition associated with abnormal cellular proliferation as compared with a healthy individual. It is also understood that "elevated transcription of mRNA" may refer to a greater amount of messenger RNA transcribed from genes the expression of which is modulated by HDAC10 either alone or in combination with other molecules.

A "host cell," as used herein, refers to a prokaryotic or eukaryotic cell that contains heterologous DNA that has been introduced into the cell by any means, e.g., electroporation, calcium phosphate precipitation, microinjection, transformation, viral infection, and the like.

"Heterologous" as used herein means "of different natural origin" or represent a non-natural state. For example, if a host cell is transformed with a DNA or gene derived from another organism, particularly from another species, that gene is heterologous with respect to that host cell and also with respect to descendants of the host cell which carry that gene. Similarly, heterologous refers to a

nucleotide sequence derived from and inserted into the same natural, original cell type, but which is present in a non-natural state, e.g. a different copy number, or under the control of different regulatory elements.

A "vector" molecule is a nucleic acid molecule into which heterologous nucleic acid may be inserted which can then be introduced into an appropriate host cell. Vectors preferably have one or more origin of replication, and one or more site into which the recombinant DNA can be inserted. Vectors often have convenient means by which cells with vectors can be selected from those without, e.g., they encode drug resistance genes. Common vectors include plasmids, viral genomes, and (primarily in yeast and bacteria) "artificial chromosomes."

"Plasmids" generally are designated herein by a lower case p preceded and/or followed by capital letters and/or numbers, in accordance with standard naming conventions that are familiar to those of skill in the art. Starting plasmids disclosed herein are either commercially available, publicly available on an unrestricted basis, or can be constructed from available plasmids by routine application of well-known, published procedures. Many plasmids and other cloning and expression vectors that can be used in accordance with the present invention are well known and readily available to those of skill in the art. Moreover, those of skill readily may construct any number of other plasmids suitable for use in the invention. The properties, construction and use of such plasmids, as well as other vectors, in the present invention will be readily apparent to those of skill from the present disclosure.

The term "isolated" means that the material is removed from its original environment (e.g., the natural environment if it is naturally occurring). For example, a naturally occurring polynucleotide or polypeptide present in a living animal is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the coexisting materials in the natural system, is isolated, even if subsequently reintroduced into the natural system. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment.

As used herein, the term "transcriptional control sequence" refers to DNA sequences, such as initiator sequences, enhancer sequences, and promoter sequences, which induce, repress, or otherwise control the transcription of protein encoding nucleic acid sequences to which they are operably linked.

As used herein, "human transcriptional control sequences" are any of those transcriptional control sequences normally found associated with the human gene encoding the novel HDAC10 polypeptide of the present invention as it is found in the respective human chromosome. It is understood that the term may also refer to transcriptional control sequences normally found associated with human genes the expression of which is modulated by HDAC10 either alone or in combination with other molecules.

As used herein, "non-human transcriptional control sequence" is any transcriptional control sequence not found in the human genome.

The term "polypeptide" is used interchangeably herein with the terms "polypeptides" and "protein(s)".

As used herein, a "chemical derivative" of a polypeptide of the invention is a polypeptide of the invention that contains additional chemical moieties not normally a part of the molecule. Such moieties may improve the molecule's solubility, absorption, biological half life, etc. The moieties may alternatively decrease the toxicity of the molecule, eliminate or attenuate any undesirable side effect of the molecule, etc. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 16th ed., Mack Publishing Co., Easton, Pa. (1980).

As used herein, "HDAC10" refers to the amino acid sequences of substantially purified HDAC10 obtained from any species, particularly mammalian, including bovine, ovine, porcine, murine, equine, and preferably human, from any source, whether natural, synthetic, semi-synthetic, or recombinant.

As used herein, "HDAC activity", including "HDAC10 activity" refers to the ability of an HDAC polypeptide to deacetylate histone proteins, including <sup>3</sup>H-labeled H4 histone peptide. Such activity may be measured according to conventional methods. A biologically "active" protein refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule.

The term "agonist", as used herein, refers to a molecule which when bound to HDAC10, causes a change in HDAC10 which modulates the activity of HDAC10. Agonists may include proteins, nucleic acids, carbohydrates, or any other molecules that bind to HDAC10.

The terms "antagonist" or "inhibitor" as used herein, refer to a molecule which when bound to HDAC10, blocks or modulates the biological activity of HDAC10. Antagonists and inhibitors may include proteins, nucleic acids, carbohydrates, or any other molecules, natural or synthetic that bind to HDAC10.

The full-length cDNA for HDAC10 is 1755 base pairs in length and it predicts a protein of 347 amino acids. The predicted HDAC10 protein possesses a putative catalytic domain which encompasses approximately 317 amino acids (~6 to 323) based upon alignments of HDAC10 with the putative catalytic domains of all of the other known HDACs. To identify the catalytic domain of HDAC10, Clustalw alignments were performed separately using HDAC10 complete peptide and catalytic domain sequences from class I HDACs (1-3 and 8) or class II HDACs (4-7).

Table 2 below shows the catalytic domain amino acids of HDACs 1-10 that align with histone deacetylase-like protein (HDLP), a bacterial protein that shares 35.2% homology with HDAC1 and possesses deacetylase activity (Finnin, M. S., Donigian, J. R., Cohen, A., Richon, V. M., Rifkind, R. A., Marks, P. A., Breslow, R., and Pavletich, N. P. (1999) *Nature* 401, 188-193).

**Table 2. HDAC catalytic amino acids**

HDAC Isoform	Amino acids in the catalytic domains of HDAC isoforms													
HDLP	Pro 22	His 131	His 132	Gly 140	Phe 141	Asp 166	Asp 168	His 170	Asp 173	Phe 198	Asp 258	Leu 265	Tyr 297	
HDAC1	Pro	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr	
HDAC2	Pro	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr	
HDAC3	Pro	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Leu	Tyr	
HDAC4	Pro	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His	
HDAC5	Pro	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His	
HDAC6-1	Pro	His	His	Gly	Tyr	Asp	Asp	His	Gln	Phe	Asp	Lys	Tyr	
HDAC6-2	Pro	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	Tyr	
HDAC7	Pro	His	His	Gly	Phe	Asp	Asp	His	Asn	Phe	Asp	Leu	His	
HDAC8	Pro	His	His	Gly	Phe	Asp	Asp	His	Asp	Phe	Asp	Met	Tyr	
HDAC 9	Pro	His	His	Gly	Phe	Asp	Asp	His	Gln	Phe	Asp	Glu	Tyr	
HDAC10	Pro 36	His 142	His 143	Gly 151	Phe 152	Asp 179	Asp 181	His 183	Asn 186	Tyr 209	Asp 261	Leu 268	Tyr 304	

Italicized amino acids represent amino acids that are not always conserved.

As a member of the HDAC family, HDAC10 may form biologically relevant complexes with proteins and display functions that have been described for other HDACs. For example, it is likely to be involved in transcription repression as a component of multi-protein complexes that often include transcription co-repressors. Thus, increased activity or expression of HDAC10 may be associated with numerous pathological conditions, including but not limited to, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis.

Thus, the identification of HDAC10 is useful for designing agents (e.g. antagonists or inhibitors) useful to ameliorate conditions associated with abnormal HDAC activity. These may include, for example, antiproliferative or antiinflammatory agents either through the use of small molecules or proteins (e.g. antibodies) directed against it or its associated proteins in the HDAC transcription repressor complexes. In addition, protein derived from the HDAC10 sequence may also be used as a therapeutic to modify host cell proliferative or inflammatory responses.

To determine the tissue distribution of HDAC10 in human, Northern analyses were performed using a blot containing mRNA isolated from various human tissues. The results indicate that overall expression level of HDAC10 is low and the highest expression level is restricted to brain, heart, skeletal muscle and kidney. Furthermore, real-time PCR experiments reveal that HDAC10 is also highly expressed in testis as well as several human cancerous cell lines. Thus, HDAC10 represents a transcribed gene.

In one aspect, the present invention relates to a novel histone deacetylase (HDAC). As outlined above, HDAC10 is clearly a member of the HDAC family since it is highly similar to other HDAC proteins, especially in the catalytic domain.

The present invention relates to an isolated polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1. For example, such a polypeptide may be a fusion protein including the amino acid sequence of the novel HDAC10. In another aspect the present invention relates to an isolated polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1, which is, in particular, the novel HDAC10.

The invention includes nucleic acid or nucleotide molecules, preferably DNA molecules, in particular encoding the novel HDAC10. Preferably, an isolated nucleic acid molecule, preferably a

DNA molecule, of the present invention encodes a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1. Likewise preferred is an isolated nucleic acid molecule, preferably a DNA molecule, encoding a polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1. Such a nucleic acid or nucleotide, in particular such a DNA molecule, preferably comprises a nucleotide sequence selected from the group consisting of (1) the nucleotide sequence as set forth in SEQ ID NO:2, which is the full-length cDNA sequence encoding the polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1; (2) the nucleotide sequence set forth in SEQ ID NO:3, which corresponds to the open reading frame of the cDNA sequence set forth in SEQ ID NO:2; (3) a nucleotide sequence capable of hybridizing under high stringency conditions to a nucleotide sequence set forth in SEQ ID NO:3; and (4) the nucleotide sequence set forth in SEQ ID NO:4, which corresponds to the endogenous genomic human DNA encoding the polypeptide consisting of the amino acid sequence set forth in SEQ ID NO:1. Such hybridization conditions may be highly stringent or less highly stringent, as described above. In instances wherein the nucleic acid molecules are deoxyoligonucleotides ("oligos"), highly stringent conditions may refer, e.g., to washing in 6X SSC/0.05% sodium pyrophosphate at 37 °C (for 14-base oligos), 48 °C (for 17-base oligos), 55 °C (for 20-base oligos), and 60 °C (for 23-base oligos). Suitable ranges of such stringency conditions for nucleic acids of varying compositions are described in Krause and Aaronson (1991), Methods in Enzymology, 200:546-556.

These nucleic acid molecules may act as target gene antisense molecules, useful, for example, in target gene regulation and/or as antisense primers in amplification reactions of target gene nucleic acid sequences. Further, such sequences may be used as part of ribozyme and/or triple helix sequences, also useful for target gene regulation. Still further, such molecules may be used as components of diagnostic methods whereby the presence of an allele causing a disease associated with abnormal HDAC10 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis, may be detected.

The invention also encompasses (a) vectors that contain at least a fragment of any of the foregoing nucleotide sequences and/or their complements (i.e., antisense); (b) vector molecules, preferably vector molecules comprising transcriptional control sequences, in particular expression vectors, that contain any of the foregoing coding sequences operatively associated with a regulatory element that directs the expression of the coding sequences; and (c) genetically engineered host cells that contain a vector molecule as mentioned herein or at least a fragment of any of the foregoing

nucleotide sequences operatively associated with a regulatory element that directs the expression of the coding sequences in the host cell. As used herein, regulatory elements include, but are not limited to, inducible and non-inducible promoters, enhancers, operators and other elements known to those skilled in the art that drive and regulate expression. Preferably, host cells can be vertebrate host cells, preferably mammalian host cells, such as human cells or rodent cells, such as CHO or BHK cells. Likewise preferred, host cells can be bacterial host cells, in particular *E.coli* cells.

Particularly preferred is a host cell, in particular of the above described type, which can be propagated in vitro and which is capable upon growth in culture of producing an HDAC10 polypeptide, in particular a polypeptide comprising or consisting of an amino acid sequence set forth in SEQ ID NO:1, wherein said cell contains some fragment or complete sequence of HDAC10 coding sequence in a construct that is controlled by one or more transcriptional control sequences that is not a transcriptional control sequence of the natural endogenous human gene encoding said polypeptide, wherein said one or more transcriptional control sequences control transcription of a DNA encoding said polypeptide. Possible transcriptional control sequences include, but are not limited to, bacterial or viral promoter sequences.

The invention includes the complete sequence of the gene as well as fragments of any of the nucleic acid sequences disclosed herein. Fragments of the nucleic acid sequences encoding the novel HDAC10 polypeptide may be used as a hybridization probe for a cDNA library to isolate other genes which have a high sequence similarity to the HDAC10 gene or similar biological activity. Probes of this type preferably have at least about 30 bases and may contain, for example, from about 30 to about 50 bases, about 50 to about 100 bases, about 100 to about 200 bases, or more than 200 bases. The probe may also be used to identify a cDNA clone that correspond to a full-length transcript and a genomic clone or clones that contain the complete HDAC10 gene including regulatory and promoter regions, exons, and introns. An example of a screen comprises isolating the coding region of the HDAC10 gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the gene of the present invention may be used to screen a library of human cDNA, genomic DNA or mRNA to determine which members of the library to which the probe hybridizes.

In addition to the gene sequences described above, homologs of such sequences, as may, for example, be present in other species, may be identified and may be readily isolated, without undue experimentation, by molecular biological techniques well known in the art. Furthermore, there may

exist genes at other genetic loci within the genome that encode proteins that have homology to one or more domains of such gene products. These genes may also be identified via similar techniques. For example, the isolated nucleotide sequence of the present invention encoding the novel HDAC10 polypeptide may be labeled and used to screen a cDNA library constructed from mRNA obtained from the organism of interest. Hybridization conditions will be of a lower stringency when the cDNA library is derived from an organism different from the type of organism from which the labeled sequence was derived. Alternatively, the labeled fragment may be used to screen a genomic library derived from the organism of interest, again, using appropriately stringent conditions. Such low stringency conditions will be well known to those of skill in the art, and will vary predictably depending on the specific organisms from which the library and the labeled sequences are derived.

Further, a previously unknown differentially expressed gene-type sequence may be isolated by performing PCR using two degenerate oligonucleotide primer pools designed on the basis of amino acid sequences within the gene of interest. The template for the reaction may be cDNA obtained by reverse transcription of mRNA prepared from human or non-human cell lines or tissue known or suspected to express a differentially expressed gene allele. The PCR product may be subcloned and sequenced to ensure that the amplified sequences represent the sequences of a differentially expressed gene-like nucleic acid sequence. The PCR fragment may then be used to isolate a complete cDNA clone by a variety of conventional methods. For example, the amplified fragment may be labeled and used to screen a bacteriophage cDNA library. Alternatively, the labeled fragment may be used to screen a genomic library.

PCR technology may also be utilized to isolate full-length cDNA sequences. For example, RNA may be isolated, following standard procedures, from an appropriate cellular or tissue source. A reverse transcription reaction may be performed on the RNA using an oligonucleotide primer specific for the most 5' end of the amplified fragment for the priming of first strand synthesis. The resulting RNA/DNA hybrid may then be "tailed" with guanines using a standard terminal transferase reaction, the hybrid may be digested with RNAase H, and second strand synthesis may then be primed with a poly-C primer. Thus, cDNA sequences upstream of the amplified fragment may easily be isolated.

In cases where the gene identified is the normal, or the wild type gene, this gene may be used to isolate mutant alleles of the gene. Isolation of mutant alleles is preferable in processes and disorders that are known or suspected to have a genetic basis. Mutant alleles may be isolated from individuals either known or suspected to have a genotype which contributes to disease symptoms

related to abnormal HDAC activity, including, but not limited to, conditions such as abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis. Mutant alleles and mutant allele products may then be used in the diagnostic assay systems described below.

A cDNA of the mutant gene may be isolated, for example, using PCR, a technique that is well known to those of skill in the art. In this case, the first cDNA strand may be synthesized by hybridizing an oligo-dT oligonucleotide to mRNA isolated from tissue known or suspected to be expressed in an individual putatively carrying the mutant allele, and by extending the new strand with reverse transcriptase. The second strand of the cDNA is then synthesized using an oligonucleotide that hybridizes specifically to the 5' end of the normal gene. Using these two primers, the product is then amplified via PCR, cloned into a suitable vector, and subjected to DNA sequence analysis through methods well known to those of skill in the art. By comparing the DNA sequence of the mutant gene to that of the normal gene, the mutation(s) responsible for the loss or alteration of function of the mutant gene product can be ascertained.

Alternatively, a genomic or cDNA library can be constructed and screened using DNA or RNA, respectively, from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. The normal gene or any suitable fragment thereof may then be labeled and used as a probe to identify the corresponding mutant allele in the library. The clone containing this gene may then be purified through methods routinely practiced in the art, and subjected to sequence analysis as described above.

Additionally, an expression library can be constructed utilizing DNA isolated from or cDNA synthesized from a tissue known to or suspected of expressing the gene of interest in an individual suspected of or known to carry the mutant allele. In this manner, gene products made by the putatively mutant tissue may be expressed and screened using standard antibody screening techniques in conjunction with antibodies raised against the normal gene product, as described below. In cases where the mutation results in an expressed gene product with altered function (e.g., as a result of a mis-sense mutation), a polyclonal set of antibodies are likely to cross-react with the mutant gene product. Library clones detected via their reaction with such labeled antibodies can be purified and subjected to sequence analysis as described above.

The present invention includes those proteins encoded by nucleotide sequences set forth in any of SEQ ID NOs:2, 3 or 4, in particular, a polypeptide that is or includes the amino acid sequence set out in SEQ ID NO:1, or fragments thereof.

Furthermore, the present invention includes proteins that represent functionally equivalent gene products. Such an equivalent differentially expressed gene product may contain deletions, additions or substitutions of amino acid residues within the amino acid sequence encoded by the differentially expressed gene sequences described, above, but which result in a silent change, thus producing a functionally equivalent differentially expressed gene product. Amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved.

Nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine. Polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine. Positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Functionally equivalent," as utilized herein, may refer to a protein or polypeptide capable of exhibiting a substantially similar *in vivo* or *in vitro* activity as the endogenous differentially expressed gene products encoded by the differentially expressed gene sequences described above. "Functionally equivalent" may also refer to proteins or polypeptides capable of interacting with other cellular or extracellular molecules in a manner substantially similar to the way in which the corresponding portion of the endogenous differentially expressed gene product would. For example, a "functionally equivalent" peptide, the sequence of which was modified from the endogenous peptide to achieve "functional equivalency, would be able, in an immunoassay, to diminish the binding of an antibody to the corresponding peptide within the endogenous protein, or the binding to the endogenous protein itself, against which the antibody was raised. An equimolar concentration of the functionally equivalent peptide will diminish the aforesaid binding of the corresponding peptide by at least about 5%, preferably between about 5% and 10%, more preferably between about 10% and 25%, even more preferably between about 25% and 50%, and most preferably between about 40% and 50%.

The polypeptides of the present invention may be produced by recombinant DNA technology using techniques well known in the art. Therefore, there is provided a method of producing a polypeptide of the present invention, which method comprises culturing a host cell having

incorporated therein an expression vector containing an exogenously-derived polynucleotide encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1 under conditions sufficient for expression of the polypeptide in the host cell, thereby causing the production of the expressed polypeptide. Optionally, said method further comprises recovering the polypeptide produced by said cell. In a preferred embodiment of such a method, said exogenously-derived polynucleotide encodes a polypeptide consisting of an amino acid sequence set forth in SEQ ID NO:1. Preferably, said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in any of SEQ ID NO:2, SEQ ID NO:3 or SEQ ID NO:4. In case of using the nucleotide sequence set forth in SEQ ID NO:3, i.e. the open reading frame, the sequence, when inserted into a vector, may be followed by one or more appropriate translation stop codons, preferably by the natural endogenous stop codon TGA beginning at nucleotide 1066 in the cDNA sequence.

Thus, methods for preparing the polypeptides and peptides of the invention by expressing nucleic acid encoding respective nucleotide sequences are described herein. Methods which are well-known to those skilled in the art can be used to construct expression vectors that contain protein coding sequences and appropriate transcriptional/translational control signals. These methods include, for example, *in vitro* recombinant DNA techniques, synthetic techniques and *in vivo* recombination/genetic recombination. Alternatively, RNA capable of encoding differentially expressed gene protein sequences may be chemically synthesized using, for example, synthesizers.

A variety of host-expression vector systems may be utilized to express the HDAC10 gene coding sequences of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, exhibit the HDAC10 gene protein of the invention *in situ*. These include, but are not limited to, microorganisms such as bacteria (e.g., *E. coli*, *B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing differentially expressed gene protein coding sequences; yeast (e.g. *Saccharomyces*, *Pichia*) transformed with recombinant yeast expression vectors containing the differentially expressed gene protein coding sequences; insect cell systems infected or transfected with recombinant virus expression vectors (e.g., baculovirus) containing the differentially expressed gene protein coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant vectors, including plasmids, (e.g., Ti plasmid) containing protein coding sequences; or mammalian cell systems (e.g. COS, CHO, BHK, 293, 3T3)

harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothioneine promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter, or the CMV promoter).

Expression of the HDAC10 of the present invention by a cell from an HDAC10 encoding gene that is native to the cell can also be performed. Such methods are known in the art. Cells that have been induced to express HDAC10 can be implanted into a desired tissue in a living animal in order to increase the local concentration of HDAC10 in the tissue.

In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the protein being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of antibodies or to screen peptide libraries, for example, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. In this respect, fusion proteins comprising hexahistidine tags may be used, such as EpiTag vectors including pCDNA3.1/His (Invitrogen, Carlsbad, CA). Other vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., 1983, EMBO J. 2:1791), in which the protein coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors; and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene protein can be released from the GST moiety. Fusion proteins containing Flag tags, such as 3X Flag (Sigma, St. Louis, MO) or myc tags, for example pCDNA3.1/myc-His (Invitrogen, Carlsbad, CA) may be used. These fusions allow coimmunoprecipitation and Western detection of proteins for which antibodies are not yet available.

Promoter regions from any desired gene can be introduced into vectors containing a reporter transcription unit, such as a chloramphenicol acetyl transferase ("CAT"), or the luciferase transcription unit, which also lack a promoter region. Restriction site or sites in the vector can be used for introducing a candidate promoter fragment; i.e., a fragment that may contain a promoter. For example, introduction into the vector of a promoter-containing fragment at the restriction site upstream of the cat gene engenders production of CAT activity, which can be detected by standard CAT assays. Vectors suitable to this end are well known and readily available. Two such vectors are

pKK232-8 and pCM7. Thus, promoters for expression of polynucleotides of the present invention include not only well-known and readily available promoters, but also promoters that readily may be obtained by the foregoing technique, using a reporter gene.

Among known bacterial promoters suitable for expression of polynucleotides and polypeptides in accordance with the present invention are the *E. coli* lacI and lacZ promoters, the T3 and T7 promoters, the T5 tac promoter, the lambda PR, PL promoters and the trp promoter. Among known eukaryotic promoters suitable in this regard are the CMV immediate early promoter, the HSV thymidine kinase promoter, the early and late SV40 promoters, the promoters of retroviral LTRs, such as those of the Rous sarcoma virus ("RSV"), and metallothionein promoters, such as the mouse metallothionein-I promoter. For example, a plasmid construct could contain a HDAC10 transcriptional control sequence fused to a reporter transcription unit that encodes the coding region of  $\beta$ -Galactosidase, chloramphenicol acetyltransferase, green fluorescent protein or luciferase. This construct could be used to screen for small molecules that modulate HDAC10 transcription. Such molecules are potential therapeutics. Furthermore, using fluorescence microscopy or Biophotonic *in vivo* imaging, a technology that produces visual and quantitative measurements in real time (Xenogen, Palo Alto, CA), expression of a fluorescent HDAC10 reporter gene could be examined to determine the effects of an HDAC10 therapeutic in mammalian cells or xenografts. Changes in these reporters in normal, diseased or drug-treated tissue or cells would be indicators of changes in HDAC10 expression or activity.

In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is one of several insect systems that can be used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter). Successful insertion of the coding sequence will result in inactivation of the polyhedrin gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedrin gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed.

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo*

recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the desired protein in infected hosts (e.g., See Logan & Shenk, 1984, Proc. Natl. Acad. Sci. USA 81:3655-3659). Specific initiation signals may also be required for efficient translation of inserted gene coding sequences. These signals include the ATG initiation codon and adjacent sequences. In cases where an entire gene, including its own initiation codon and adjacent sequences, is inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only a portion of the gene coding sequence is inserted, exogenous translational control signals, including, perhaps, the ATG initiation codon, must be provided. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc.. Other common systems are based on SV40, retrovirus or adeno-associated virus. Selection of appropriate vectors and promoters for expression in a host cell is a well known procedure and the requisite techniques for expression vector construction, introduction of the vector into the host and expression in the host per se are routine skills in the art. Generally, recombinant expression vectors will include origins of replication, a promoter derived from a highly expressed gene to direct transcription of a downstream structural sequence, and a selectable marker to permit isolation of vector containing cells after exposure to the vector.

In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells that possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERO, BHK, HeLa, COS, MDCK, 293, 3T3, WI38, etc. and are well known to one of skill in the art.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines that stably express a differentially expressed protein product of a gene may be

engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines that express the differentially expressed gene protein. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that affect the endogenous activity of the expressed protein.

A number of selection systems may be used, including but not limited to, the herpes simplex virus thymidine kinase, hypoxanthine-guanine phosphoribosyltransferase and adenine phosphoribosyltransferase genes can be employed in tk<sup>-</sup>, hgprt<sup>-</sup> or aprt<sup>-</sup> cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for dhfr, which confers resistance to methotrexate, gpt, which confers resistance to mycophenolic acid; neo, which confers resistance to the aminoglycoside G-418; and hygro, which confers resistance to hygromycin genes.

An alternative fusion protein system allows for the ready purification of non-denatured fusion proteins expressed in human cell lines. In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the gene's open reading frame is translationally fused to an amino-terminal tag consisting of six histidine residues. Extracts from cells infected with recombinant vaccinia virus are loaded onto Ni<sup>2+</sup> nitriloacetic acid-agarose columns and histidine-tagged proteins are selectively eluted with imidazole-containing buffers.

When used as a component in assay systems such as those described below, a protein of the present invention may be labeled, either directly or indirectly, to facilitate detection of a complex formed between the protein and a test substance. Any of a variety of suitable labeling systems may be used including, but not limited to, radioisotopes such as <sup>125</sup>I; enzyme labeling systems that generate a detectable calorimetric signal or light when exposed to substrate; and fluorescent labels.

Where recombinant DNA technology is used to produce a protein of the present invention for such assay systems, it may be advantageous to engineer fusion proteins that can facilitate labeling, immobilization, detection and/or isolation

Indirect labeling involves the use of a protein, such as a labeled antibody, which specifically binds to a polypeptide of the present invention. Such antibodies include but are not limited to polyclonal, monoclonal, chimeric, single chain, Fab fragments and fragments produced by an Fab expression library.

In another embodiment, nucleic acids comprising a sequence encoding HDAC10 protein or functional derivative thereof, may be administered to promote normal biological function, for example, normal transcriptional regulation, by way of gene therapy. Gene therapy refers to therapy performed by the administration of a nucleic acid to a subject. In this embodiment of the invention, the nucleic acid produces its encoded protein that mediates a therapeutic effect by promoting normal transcriptional regulation.

Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

In a preferred aspect, the therapeutic comprises a HDAC10 nucleic acid that is part of an expression vector that expresses a HDAC10 protein or fragment or chimeric protein thereof in a suitable host. In particular, such a nucleic acid has a promoter operably linked to the HDAC10 coding region, said promoter being inducible or constitutive, and, optionally, tissue-specific. In another particular embodiment, a nucleic acid molecule is used in which the HDAC10 coding sequences and any other desired sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the HDAC10 nucleic acid.

Delivery of the nucleic acid into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vector, or indirect, in which case, cells are first transformed with the nucleic acid *in vitro*, then transplanted into the patient. These two approaches are known, respectively, as *in vivo* or *ex vivo* gene therapy.

In a specific embodiment, the nucleic acid is directly administered *in vivo*, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, for example, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by infection using a defective or

attenuated retroviral or other viral vector, or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering it in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, a nucleic acid-ligand complex can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted *in vivo* for cell specific uptake and expression, by targeting a specific receptor. Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

In a specific embodiment, a viral vector that contains the HDAC10 nucleic acid is used. For example, a retroviral vector can be used. These retroviral vectors have been modified to delete retroviral sequences that are not necessary for packaging of the viral genome and integration into host cell DNA. The HDAC10 nucleic acid to be used in gene therapy is cloned into the vector, which facilitates delivery of the gene into a patient.

Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Adeno-associated virus (AAV) has also been proposed for use in gene therapy.

Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

In this embodiment, the nucleic acid is introduced into a cell prior to administration *in vivo* of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or

- 25 -

bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

The resulting recombinant cells can be delivered to a patient by various methods known in the art. In a preferred embodiment, epithelial cells are injected, e.g., subcutaneously. In another embodiment, recombinant skin cells may be applied as a skin graft onto the patient. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, available cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as T lymphocytes, B lymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

In an embodiment in which recombinant cells are used in gene therapy, a HDAC10 nucleic acid is introduced into the cells such that it is expressible by the cells or their progeny, and the recombinant cells are then administered *in vivo* for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem-and/or progenitor cells that can be isolated and maintained *in vitro* can potentially be used in accordance with this embodiment of the present invention. Such stem cells include but are not limited to hematopoietic stem cells (HSC), stem cells of epithelial tissues such as the skin and the lining of the gut, embryonic heart muscle cells, liver stem cells, and neural stem cells.

Epithelial stem cells (ESCs) or keratinocytes can be obtained from tissues such as the skin and the lining of the gut by known procedures. In stratified epithelial tissue such as the skin, renewal occurs by mitosis of stem cells within the germinal layer, the layer closest to the basal lamina. Stem cells within the lining of the gut provide for a rapid renewal rate of this tissue. ESCs or keratinocytes obtained from the skin or lining of the gut of a patient or donor can be grown in tissue culture. If the ESCs are provided by a donor, a method for suppression of host versus graft reactivity (e.g., irradiation, drug or antibody administration to promote moderate immunosuppression) can also be used.

With respect to hematopoietic stem cells (HSC), any technique which provides for the isolation, propagation, and maintenance in vitro of HSC can be used in this embodiment of the invention. Techniques by which this may be accomplished include (a) the isolation and establishment of HSC cultures from bone marrow cells isolated from the future host, or a donor, or (b) the use of previously established long-term HSC cultures, which may be allogeneic or xenogeneic. Non-autologous HSC are used preferably in conjunction with a method of suppressing transplantation immune reactions of the future host/patient. In a particular embodiment of the present invention, human bone marrow cells can be obtained from the posterior iliac crest by needle aspiration. In a preferred embodiment of the present invention, the HSCs can be made highly enriched or in substantially pure form. This enrichment can be accomplished before, during, or after long-term culturing, and can be done by any techniques known in the art. Long-term cultures of bone marrow cells can be established and maintained by using, for example, modified Dexter cell culture techniques or Witlock-Witte culture techniques.

In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription.

A further embodiment of the present invention relates to a purified antibody or a fragment thereof which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1 or to a fragment of said polypeptide. A preferred embodiment relates to a fragment of such an antibody, which fragment is an Fab or F(ab')<sub>2</sub> fragment. In particular, the antibody can be a polyclonal antibody or a monoclonal antibody.

Methods for the production of antibodies capable of specifically recognizing one or more differentially expressed gene epitopes are known to one of ordinary skill in the art. Such antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab')<sub>2</sub> fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above. Such antibodies may be used, for example, in the detection of a fingerprint, target, gene in a biological sample, or, alternatively, as a method for the inhibition of abnormal target gene activity. Thus, such antibodies may be utilized as part of disease treatment methods, and/or may be used as part of diagnostic techniques whereby patients may be tested for abnormal levels of the HDAC10 polypeptide, or for the presence of abnormal forms of the HDAC10 polypeptide.

For the production of antibodies to the HDAC10 polypeptide, various host animals may be immunized by injection with the HDAC10 polypeptide, or a portion thereof. Such host animals may include but are not limited to rabbits, mice, and rats, to name but a few. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and Corynebacterium parvum.

Polyclonal antibodies are heterogeneous populations of antibody molecules derived from the sera of animals immunized with an antigen, such as target gene product, or an antigenic functional derivative thereof. For the production of polyclonal antibodies, host animals such as those described above, may be immunized by injection with the HDAC10 polypeptide, or a portion thereof, supplemented with adjuvants as also described above.

Monoclonal antibodies, which are homogeneous populations of antibodies to a particular antigen, may be obtained by any technique that provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

In addition, techniques developed for the production of "chimeric antibodies" by splicing the genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable or hypervariable region derived from a murine mAb and a human immunoglobulin constant region.

Alternatively, techniques described for the production of single chain antibodies can be adapted to produce differentially expressed gene-single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide.

Most preferably, techniques useful for the production of "humanized antibodies" can be adapted to produce antibodies to the polypeptides, fragments, derivatives, and functional equivalents disclosed herein.

Antibody fragments that recognize specific epitopes may be generated by known techniques. For example, such fragments include but are not limited to: the F(ab')<sub>2</sub> fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')<sub>2</sub> fragments. Alternatively, Fab expression libraries may be constructed (Huse et al., 1989, Science, 246:1275-1281) to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity.

An antibody of the present invention can be preferably used in a method for the diagnosis of a condition associated with abnormal HDAC10 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis, in a human which comprises: measuring the amount of a polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1, or fragments thereof, in an appropriate tissue or cell from a human suffering from a condition associated with abnormal HDAC10 activity, wherein the presence of an elevated amount of said polypeptide or fragments thereof, relative to the amount of said polypeptide or fragments thereof in the respective tissue from a human not suffering from a condition associated with abnormal HDAC10 activity is diagnostic of said human's suffering from such condition. Such a method forms a further embodiment of the present invention. Preferably, said detecting step comprises contacting said appropriate tissue or cell with an antibody which

specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1 or a fragment thereof and detecting specific binding of said antibody with a polypeptide in said appropriate tissue or cell, wherein detection of specific binding to a polypeptide indicates the presence of a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1 or a fragment thereof.

Particularly preferred, for ease of detection, is the sandwich assay, of which a number of variations exist, all of which are intended to be encompassed by the present invention.

For example, in a typical forward assay, unlabeled antibody is immobilized on a solid substrate and the sample to be tested brought into contact with the bound molecule. After a suitable period of incubation time sufficient to allow formation of an antibody-antigen binary complex, a second antibody, labeled with a reporter molecule capable of inducing a detectable signal, is then added and incubated, allowing time sufficient for the formation of a ternary complex of antibody-antigen-labeled antibody. Any unreacted material is washed away, and the presence of the antigen is determined by observation of a signal, or may be quantitated by comparing with a control sample containing known amounts of antigen. Variations on the forward assay include the simultaneous assay, in which both sample and antibody are added simultaneously to the bound antibody, or a reverse assay in which the labeled antibody and sample to be tested are first combined, incubated and added to the unlabeled surface bound antibody. These techniques are well known to those skilled in the art, and the possibility of minor variations will be readily apparent. As used herein, "sandwich assay" is intended to encompass all variations on the basic two-site technique. For the immunoassays of the present invention, the only limiting factor is that the labeled antibody be an antibody that is specific for the HDAC10 polypeptide or a fragment thereof.

The most commonly used reporter molecules in this type of assay are either enzymes, fluorophore- or radionuclide-containing molecules. In the case of an enzyme immunoassay an enzyme is conjugated to the second antibody, usually by means of glutaraldehyde or periodate. As will be readily recognized, however, a wide variety of different ligation techniques exist, which are well known to the skilled artisan. Commonly used enzymes include horseradish peroxidase, glucose oxidase, beta-galactosidase and alkaline phosphatase, among others. The substrates to be used with the specific enzymes are generally chosen for the production, upon hydrolysis by the corresponding enzyme, of a detectable color change. For example, p-nitrophenyl phosphate is suitable for use with alkaline phosphatase conjugates; for peroxidase conjugates, 1,2-phenylenediamine or toluidine are

commonly used. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. A solution containing the appropriate substrate is then added to the tertiary complex. The substrate reacts with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated, usually spectrophotometrically, to give an evaluation of the amount of HDAC10 which is present in the serum sample.

Alternately, fluorescent compounds, such as fluorescein and rhodamine, may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labeled antibody absorbs the light energy, inducing a state of excitability in the molecule, followed by emission of the light at a characteristic longer wavelength. The emission appears as a characteristic color visually detectable with a light microscope. Immunofluorescence and EIA techniques are both very well established in the art and are particularly preferred for the present method. However, other reporter molecules, such as radioisotopes, chemiluminescent or bioluminescent molecules may also be employed. It will be readily apparent to the skilled artisan how to vary the procedure to suit the required use.

This invention also relates to the use of polynucleotides of the present invention as diagnostic reagents. In particular, the invention relates to a method for the diagnosis of a condition associated with abnormal HDAC10 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis in a human which comprises: detecting elevated transcription of messenger RNA transcribed from the natural endogeneous human gene encoding the polypeptide consisting of an amino acid sequence set forth in SEQ ID NO:1 in an appropriate tissue or cell from a human, wherein said elevated transcription is diagnostic of said human's suffering from the condition associated with abnormal HDAC10 expression or activity. In particular, said natural endogeneous human gene comprises the nucleotide sequence set forth in SEQ ID NO:4. In a preferred embodiment such a method comprises contacting a sample of said appropriate tissue or cell or contacting an isolated RNA or DNA molecule derived from that tissue or cell with an isolated nucleotide sequence of at least about 20 nucleotides in length that hybridizes under high stringency conditions with the isolated nucleotide sequence encoding a polypeptide consisting of an amino acid sequence set forth in SEQ ID NO:1.

Detection of a mutated form of the gene characterized by the polynucleotide of SEQ ID NO:4 which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a

diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals carrying mutations in the gene may be detected at the DNA level by a variety of techniques.

Nucleic acids, in particular mRNA, for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR or other amplification techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled nucleotide sequences which encode the HDAC10 polypeptide of the present invention. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence differences may also be detected by alterations in electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing. Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method. In another embodiment, an array of oligonucleotides probes comprising nucleotide sequence encoding the HDAC10 polypeptide of the present invention or fragments of such a nucleotide sequence can be constructed to conduct efficient screening of e.g., genetic mutations. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability.

The diagnostic assays offer a process for diagnosing or determining a susceptibility to disease through detection of mutation in the HDAC10 gene by the methods described. In addition, such diseases may be diagnosed by methods comprising determining from a sample derived from a subject an abnormally decreased or increased level of polypeptide or mRNA. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example, nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radioimmunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

Thus in another aspect, the present invention relates to a diagnostic kit which comprises:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO:2, 3 or 4, or a fragment thereof;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:1 or a fragment thereof; or
- (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:1.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly to a disease or condition associated with abnormal HDAC10 expression or activity, for example, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel disease, host inflammatory or immune response, or psoriasis.

The nucleotide sequences of the present invention are also valuable for chromosome localization. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene-associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (coinheritance of physically adjacent genes).

The differences in the cDNA or genomic sequence between affected and unaffected individuals can also be determined. If a mutation is observed in some or all of the affected individuals but not in any normal individuals, then the mutation is likely to be the causative agent of the disease.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition, in conjunction with a pharmaceutically acceptable carrier, excipient or diluent, for any of the therapeutic effects discussed above. Such pharmaceutical compositions may consist of HDAC10, antibodies to that polypeptide, mimetics, agonists, antagonists, or inhibitors of HDAC10 function. The compositions may be administered alone or in combination with at least one other agent, such as stabilizing compound, which may be administered in any sterile, biocompatible pharmaceutical carrier, including, but not limited to, saline, buffered saline, dextrose, and water. The

compositions may be administered to a patient alone, or in combination with other agents, drugs or hormones.

In addition, any of the therapeutic proteins, antagonists, antibodies, agonists, antisense sequences or vectors described above may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects. Antagonists and agonists of HDAC10 may be made using methods that are generally known in the art.

The pharmaceutical compositions encompassed by the invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-articular, intra-arterial, intramedullary, intrathecal, intraventricular, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

In addition to the active ingredients, these pharmaceutical compositions may contain suitable pharmaceutically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, Pa.).

Pharmaceutical compositions for oral administration can be formulated using pharmaceutically acceptable carriers well known in the art in dosages suitable for oral administration. Such carriers enable the pharmaceutical compositions to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions, and the like, for ingestion by the patient.

Pharmaceutical preparations for oral use can be obtained through combination of active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are carbohydrate or protein fillers, such as sugars, including lactose, sucrose, mannitol, or sorbitol; starch from corn, wheat, rice, potato, or other plants; cellulose, such as methyl cellulose, hydroxypropylmethyl-cellulose, or sodium carboxymethylcellulose; gums including arabic and

tragacanth; and proteins such as gelatin and collagen. If desired, disintegrating or solubilizing agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, alginic acid, or a salt thereof, such as sodium alginate.

Dragee cores may be used in conjunction with suitable coatings, such as concentrated sugar solutions, which may also contain gum arabic, talc, polyvinylpyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for product identification or to characterize the quantity of active compound, i.e., dosage.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a coating, such as glycerol or sorbitol. Push-fit capsules can contain active ingredients mixed with a filler or binders, such as lactose or starches, lubricants, such as talc or magnesium stearate, and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid, or liquid polyethylene glycol with or without stabilizers.

Pharmaceutical formulations suitable for parenteral administration may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiologically buffered saline. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Non-lipid polycationic amino polymers may also be used for delivery. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

For topical or nasal administration, penetrants appropriate to the particular barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

The pharmaceutical compositions of the present invention may be manufactured in a manner that is known in the art, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping, or lyophilizing processes.

The pharmaceutical composition may be provided as a salt and can be formed with many acids, including but not limited to, hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents than are the corresponding free base forms. In other cases, the preferred preparation may be a lyophilized powder which may contain any or all of the following: 1-50 mM histidine, 0. 1%-2% sucrose, and 2-7% mannitol, at a pH range of 4.5 to 5.5, that is combined with buffer prior to use.

After pharmaceutical compositions have been prepared, they can be placed in an appropriate container and labeled for treatment of an indicated condition. For administration of the HDAC10, such labeling would include amount, frequency, and method of administration.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models, usually mice, rabbits, dogs, or pigs. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example HDAC10 or fragments thereof, antibodies of HDAC10, agonists, antagonists or inhibitors of HDAC10, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., ED50 (the dose therapeutically effective in 50% of the population) and LD50 (the dose lethal to 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index, and it can be expressed as the ratio, LD50/ED50. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies is used in formulating a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject that requires treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, general health of the subject, age, weight, and gender of the subject, diet, time and frequency of administration, drug combination(s), reaction sensitivities, and tolerance/response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or once every two weeks depending on half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from 0.1 to 100,000 micrograms, up to a total dose of about 1 g, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc. Pharmaceutical formulations suitable for oral administration of proteins are known in the art.

All patent applications, patents and literature references cited herein are hereby incorporated by reference in their entirety.

The following Examples illustrate the present invention, without in any way limiting the scope thereof.

**Example 1: HDAC10 protein expression *in vivo***

An expression vector containing HDAC10's coding sequences plus the Flag-epitope encoding sequences at the C-terminus is transfected into 293 embryonic kidney cells using the GenePORTER2 transfection reagent (Gene Therapy System Inc., San Diego, CA). Forty-eight hr. after transfection, cell lysates are prepared from the transfected cells and 10 µg of total protein is subjected to SDS-PAGE on a 10% Tris-glycine gel. The proteins are then transferred onto a PVDF membrane and probed with an anti-Flag antibody, followed by a secondary antibody that is conjugated with horseredish peroxidase, which allows for detection of signal using enhanced luminescence reagents. The anti-Flag antibody detects the HDAC10-Flag fusion protein as a single band of 39 kDa in size, which agrees with the estimated size of HDAC10 protein based on its amino acid composition.

Example 2: Distribution of HDAC10 mRNA in normal human tissues and cancer cell lines

A multiple human tissue Northern blot is purchased from Clontech (Palo Alto, CA). A <sup>32</sup>P-labeled probe corresponding to HDAC10 cDNA (nucleotide no.181 to no.1122) is prepared using the Rediprime DNA labeling system (Amersham Pharmacia Biotech) according to the manufacturer's instructions. The Northern blot is pre-hybridized and hybridized in the presence of the <sup>32</sup>P-labeled probe under stringent conditions according to the manufacturer's protocol. A probe corresponding to human actin cDNA (Clontech) is used as a control for the relative amount of mRNA in each lane. Results of Northern analyses indicate that there are two spliced variant forms of HDAC10 mRNA, one is ~1.7kb, which agrees with the size of the full-length cDNA (SEQ ID NO:2); the other is ~3.2kb and is expressed at a higher level. The larger transcript agrees with the size of a *Macac fascicularis* brain cDNA clone (GenBank™ accession #AB052134), which encodes a truncated HDAC10 polypeptide (minus the first 29 amino acids) with 3 conservative amino acid substitutions. Northern analyses also show that overall expression level of HDAC10 mRNA is low and high expression level is restricted to brain, heart, skeletal muscle and kidney. These findings imply that the HDAC10 gene is expressed in normal human tissues and that HDAC10's function may be tissue-specific.

In addition to Northern blotting, the Real-time PCR technique is used to examine HDAC10 mRNA distribution in normal human tissues as well as several human cancer cell lines. These experiments confirm findings of the Northern analyses; in addition, they reveal high expression level of HDAC10 in testis. Furthermore, our data indicate that large amount of HDAC10 mRNA is also found in a non-small cell lung carcinoma cell line, a rhabdomyosarcoma muscle tumor line, a urinary bladder cancer cell line and an osteosarcoma cell line. Taken together, these results indicate that HDAC10 may function not only in normal human tissues, but also in the development and/or maintenance of human cancers.

Example 3: *In vitro* HDAC enzyme assay

To determine whether the putative HDAC "10" is an active deacetylase, transfected Flag epitope-tagged recombinant HDAC10 is used to measure the ability of HDAC10 to deacetylate histone H4 peptide. Enzymatic activity may be determined according to conventional methods, such as the following techniques:

*Preparation of HDAC10-Flag expression vector.* Using conventional techniques in molecular biology, a Flag-epitope sequence is added to the C-terminus of HDAC10 coding sequences (SEQ ID NO:3) by PCR. The PCR primers are:

Forward: 5'-GAGGATCCACCATGCTACACACAACCCAGCTG- 3'

Reverse: 5'-GCGTCTAGACTACTTGTCATCGTCGTCCTGTAATCAGCCCGGGG-  
ACTGCAGGGGGAAAG- 3'.

The BamHI and XbaI restriction enzyme cutting sites are underlined, the ATG translational start site is bolded in the forward primer and the Flag-epitope encoding sequences are bolded in the reverse primer. The Flag-tagged HDAC10 PCR fragment is cloned into the pcDNA3.1(+) expression vector between the BamHI and XbaI sites.

*Transfection and Immunoprecipitation.* Approximately  $1 \times 10^7$  293 human embryonic kidney cells were grown in a 15-cm<sup>2</sup> plate (~50% confluent) on the day of transfection. GenePORTER transfection reagent (Gene Therapy Systems, Inc., San Diego, CA) is used to transfect 30 µg of plasmid DNA per plate of cells according to manufacturer's instructions. Forty-eight hr after transfection, cells are washed twice with ice-cold phosphate-buffered saline (PBS) and resuspended in 1 mL ice-cold lysis buffer (50 mM Tris-Cl, pH 7.4, 120 mM NaCl, 0.5 mM EDTA, 0.5% NP-40) supplemented with EDTA-free protease inhibitor complete (Roche Molecular Biochemicals, Indianapolis, IN). The lysate is incubated at 4°C for 20 min on a rotator, followed by spinning at 12,000 x g for 20 min at 4°C. The soluble supernatant is collected and used for immunoprecipitation with 20 µl anti-FLAG M2 affinity gel (Sigma, Saint Louis, MO) at 4°C overnight. As a negative control, 1 mL lysis buffer is used instead of the cell lysate. The immunoprecipitated complex is pelleted by centrifugation and washed three times with 1 mL ice-cold lysis buffer, four times with lysis buffer containing 1 M NaCl and three times with 1 mL HDAC assay buffer (10 mM Tris-Cl, pH 8.0, 10 mM NaCl, 10% glycerol).

*In vitro HDAC enzyme assay.* The immunoprecipitated complex is suspended in 30 µl HDAC assay buffer containing 30,000 cpm of the acetylated histone H4 peptide. Histone deacetylase activity is determined after incubation at 37°C for 3 hr as described (Emiliani, S., Fischle, W., Van Lint, C., Al-Abed, Y., and Verdin, E. (1998) *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2795-2800).

Results of the *in vitro* HDAC enzyme assays show that cells expressing the HDAC10-Flag fusion protein contain 2.5-3 fold higher enzyme activity than cells expressing the pcDNA3.1(+) vector alone. Therefore, HDAC10 is likely to contain intrinsic histone deacetylase enzyme activity.

**Example 4: Identification of HDAC10 associated protein**

Using conventional methods, proteins in the same complex as HDAC10 may be identified by their ability to coimmunoprecipitate with HDAC10-Flag fusion protein. The HDAC10-Flag expression vector or the vector alone is transfected into 293 cells and cell lysates are prepared as described above. The lysates are precleared with Sepharose A/G plus agarose beads, followed by immunoprecipitation using anti-Flag antibody at 4°C overnight on a rotator as described in example 3. The immune complexes are washed twice with ice-cold lysis buffer (see example 3), twice with lysis buffer containing 1 M NaCl and twice with PBS. The final complexes are separated by SDS-PAGE on 10% Tris-glycine gels, transferred onto a PVDF membrane and probed with antibodies against known HDAC-associated proteins or other HDACs. Conversely, the immunoprecipitation could be done using antibodies of choice, and the resulting immune complexes could be probed with anti-Flag antibody.

What is claimed is:

1. An isolated polypeptide comprising the amino acid sequence set forth in SEQ ID NO:1.
2. An isolated DNA comprising a nucleic acid sequence that encodes the polypeptide of claim 1.
3. A vector molecule comprising at least a fragment of the isolated DNA according to claim 2.
4. The vector molecule according to claim 3 comprising transcriptional control sequences.
5. A host cell comprising the vector molecule according to claim 4.
6. The isolated DNA according to claim 2, comprising a nucleotide sequence selected from the group consisting of (1) the nucleotide sequence set forth in SEQ ID NO:2; (2) the nucleotide sequence set forth in SEQ ID NO:3; (3) a nucleotide sequence capable of hybridizing under high stringency conditions to a nucleotide sequence set forth in SEQ ID NO:3; and (4) the nucleotide sequence set forth in SEQ ID NO:4.
7. A vector molecule comprising the isolated DNA molecule according to claim 6, or a fragment thereof.
8. The vector molecule according to claim 7 comprising transcriptional control sequences.
9. A host cell comprising the vector molecule according to claim 8.
10. A host cell which can be propagated *in vitro* and which is capable upon growth in culture of expressing HDAC 10, wherein said cell comprises at least one transcriptional control sequence that is not a transcriptional control sequence of the natural endogenous human gene encoding HDAC 10, wherein said one or more transcriptional control sequences control transcription of a DNA encoding HDAC 10.
11. A method for the diagnosis of a condition associated with abnormal regulation of gene expression which includes, abnormal cell proliferation, cancer, atherosclerosis, inflammatory bowel

disease, host inflammatory or immune response, or psoriasis in a human which comprises: detecting abnormal transcription of messenger RNA transcribed from the natural endogenous human gene encoding HDAC 10 in an appropriate tissue or cell from a human, wherein said abnormal transcription is diagnostic of said condition.

12. The method of claim 11, wherein said natural endogenous human gene comprises the nucleotide sequence set forth in SEQ ID NO:4.

13. A method for the diagnosis of a condition associated with abnormal HDAC10 expression or activity in a human which comprises:

measuring the amount of HDAC 10, or fragments thereof, in an appropriate tissue or cell from a human suffering from said condition wherein the presence of an abnormal amount of said polypeptide or fragments thereof, relative to the amount of said polypeptide or fragments thereof in the respective tissue from a human not suffering from said condition associated with abnormal HDAC10 expression or activity is diagnostic of said human's suffering from said condition.

14. The method of claim 13, wherein said detecting step comprises contacting said appropriate tissue or cell with an antibody which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1 or a fragment thereof and detecting specific binding of said antibody with a polypeptide in said appropriate tissue or cell, wherein detection of specific binding to a polypeptide indicates the presence of a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1 or a fragment thereof.

15. An antibody or a fragment thereof which specifically binds to a polypeptide that comprises the amino acid sequence set forth in SEQ ID NO:1 or to a fragment of said polypeptide.

16. An antibody fragment according to claim 15 which is an Fab or F(ab')<sub>2</sub> fragment.

17. An antibody according to claim 15 which is a polyclonal antibody.

18. An antibody according to claim 15 which is a monoclonal antibody.

19. A method for producing an HDAC 10 polypeptide, which method comprises:

culturing a host cell having incorporated therein an expression vector comprising an exogenously-derived polynucleotide encoding a polypeptide comprising an amino acid sequence as set forth in SEQ ID NO:1 or a nucleotide sequence capable of hybridizing under high stringency conditions to a complement of said polynucleotide, under conditions sufficient for expression of the polypeptide in the host cell, thereby causing the production of the expressed polypeptide.

20. The method according to claim 19, wherein said exogenously-derived polynucleotide hybridizes under stringent conditions to the nucleotide sequence as set forth in SEQ ID NO:2.
21. The method according to claim 19, wherein said exogenously-derived polynucleotide comprises the nucleotide sequence as set forth in SEQ ID NO:3.
22. A histone deacetylase which comprises the catalytic domain of HDAC 10.

- 44 -

## SEQ ID NO:1

MLHTTQLYQH VPETPWPIVY SPRYNITFMG LEKLHPFDAG KWGKVINFLK EEKLLSDSML 60  
 VEAREASEED LLVVHTRRYL NELKWSFAVA TITEIPPVIF LPNFLVQRKV LRPLRTQTGG 120  
 TIMAGKLAVE RGWAINVGGS FHHCSSDRGG GFCAVADITL AIKFLFERVE GISRATIIDL 180  
 DAHQGNNGHER DFMDDKRVYI MDVYNRHYP GDRFAKQAIR RKVELEGTE DDEYLDKVER 240  
 NIKKSLQEHL PDVVVYNAGT DILEGDRLLGG LSISPAGIVK RDELVFRMVR GRRVPILMVT 300  
 SGGYQKRTAR IIADSILNLF GLGLIGPESP SVSAQNSDTP LLPPAVP

## SEQ ID NO:2

1 agctttggga gggccggccc cgggatgcta cacacaaccc agctgtacca gcatgtgcc  
 61 gagacacccct ggccaatcggt gtactcgccg cgctacaaca tcacatttcataatggggcctggag  
 121 aagctgcata cctttgatgc cggaaaatgg ggcaaagtga tcaatttcataatggggcctggag  
 181 aagcttctgt ctgacagcat gctgggtggag ggcggggagg cctcggagga ggacctgctg  
 241 gtgggtgcaca cgaggcgcta tcttaatggat ctcaagtgggt ctttgcgtgt tgctaccatc  
 301 acagaaatcc ccccccgttat cttccctcccc aacttcccttg tgcagaggaa ggtgctgagg  
 361 ccccttcgga cccagacagg aggaaccata atggccggggaa agctggctgt ggagcgaggc  
 421 tggggccatca acgtgggggg tggcttccac cactgctcca ggcggccatc cgggggcttc  
 481 tgcctatcgccatc aagtttctgt ttgagcgtgt ggagggcatc  
 541 tccagggtataatggat ccatacattga tcttgcattgc catcaggggca atgggcattga gcgagacttc  
 601 atggacgaca agcgtgtgtat catcatggat gtctacaacc gccacatcta cccaggggac  
 661 cgctttggcca agcaggccat caggcgaaag gtggagctgg agtggggcac agaggatgt  
 721 gagttacctgg ataagggtggat gaggaacatc aagaatccc tccaggagca cctggccgac  
 781 gtgggtgtat acaatgcagg caccgcacatc ctcggggggg accgccttgg ggggctgtcc  
 841 atcagccccag cgggcacatcgat gaaaggat gagctgggtgt tccggatgggt ccgtggccgc  
 901 cgggtggccca tccttgcgtat gacccatc gggatccatc aagggatgttgcgtatgg  
 961 gctgactcca tactaatcttgcgtat gtttgcgtat gggatccatc acccagcgatc  
 1021 tccgcacaga acttagacac accgcgttgcgtat cccctgcac tggccctgacc cttgtgc  
 1081 tgcctgttcac gtggccctgc ctatccgc cttatgcgtt tttgttttcatgg  
 1141 ggtgggtggag gcagccatca gtggatggggccatc aatccctgg ctggggcctgg  
 1201 gagctggccc ttccctact tttccctgtt ggaaggccaga agggcttggat gcctctatgg  
 1261 gtggggccatca aaggcagacccatca ctgtgtccca gggggacccatca cacaatggatc  
 1321 ggtccaggatca ggcaggcacttgcgtat taactgagaa ttggagggat caggatggatcc  
 1381 gcgaggccc tggcttggggat gtttgcgtat tttgagaacg gcaatccatc gtcggatgt  
 1441 ggaagcttcc acctccatcc tggatggccatca tggatccatc tggggcctcc  
 1501 ttggatcatgg gatttgcgtatcc cctttgcgtatcc caatggatggat gggatccatc  
 1561 ggtggccatca ggagggtggat gggatccatc gggatccatc  
 1621 tggggaaacctt gggccatggat gtttgcgtatcc tggatccatc  
 1681 tctggatgttcc cccctcaata aagcaaggatc tggatccatc  
 1741 aaaaaaaaaaaaaaaa

## SEQ ID NO:3

25 atgcta cacacaaccc agctgtacca gcatgtgcc  
 61 gagacacccct ggccaatcggt gtactcgccg cgctacaaca tcacatttcataatggggcctggag  
 121 aagctgcata cctttgatgc cggaaaatgg ggcaaagtga tcaatttcataatggggcctggag  
 181 aagcttctgt ctgacagcat gctgggtggag ggcggggagg cctcggagga ggacctgctg  
 241 gtgggtgcaca cgaggcgcta tcttaatggat ctcaagtgggt ctttgcgtgt tgctaccatc  
 301 acagaaatcc ccccccgttat cttccctcccc aacttcccttg tgcagaggaa ggtgctgagg  
 361 ccccttcgga cccagacagg aggaaccata atggccggggaa agctggctgt ggagcgaggc  
 421 tggggccatca acgtgggggg tggcttccac cactgctcca ggcggccatc cgggggcttc  
 481 tgcctatcgccatc aagtttctgt ttgagcgtgt ggagggcatc  
 541 tccagggtataatggat ccatacattga tcttgcattgc catcaggggca atgggcattga gcgagacttc  
 601 atggacgaca agcgtgtgtat catcatggat gtctacaacc gccacatcta cccaggggac  
 661 cgctttggcca agcaggccat caggcgaaag gtggagctgg agtggggcac agaggatgt

721 gagtacctgg ataagggtgga gaggAACATC aAGAAATCCC TCCAGGAGCA CCGCCGAC  
781 gtgggttat acaatgcagg caccgacatc ctcgaggggg accgccttgg ggggtgtcc  
841 atcagccccag cgggcatcgtaa gaagcgggat gagctgggtt tccggatggt ccgtggccgc  
901 cggtgccccaa tccttatggt gacccatggc gggtaaccaga agcgcacagc ccccatcatt  
961 gctgactcca tacttaatct gtttggcctg gggctcattt ggcctgagtc acccagcgtc  
1021 tccgcacaga actcagacac accgctgtttt cccccctgcag tgccccatgtt

SEQ ID NO:4

ctcacacaaa cccagctgt a ccagcatgtg ccagagacac gctggccaat cggtactcg 17663325  
ccgcgtaca acatcacctt catgggcctg gagaagctgc atcccatttg tgccggaaaa 17663265  
tggggcaaaag tgcattttt cctaaaaggat atggaaaggtc ccccttggac tctcatctgc 17663205  
ttcctccaaac ccacctgtcc tctccgtctt catccccaaac ataaggctca ggctctctcc 17663145  
catcttcaagt ttcaagccctc ggatggccctt ccaccatgc ttccggccaa aatgatttt 17663085  
ccaacacaga ctccataatca cgatatgtat tccctgactc agactctccc tggctccca 17663025  
tcctgtggc ctaagtctgt cctctgcccc agaggcttag tggaaaggta gctgattact 17662965  
gatgggcaca gggaaaggta agcttggagg agtcatttc ctaaggttca gagactcagg 17662905  
aggtagagca cctccacccg acettcttgc attacagat ggggaaattt tgccttagaa 17662845  
tgatttagaa acatgtgcac ccaattccag tccagtcctc acagcaggcc tcgggtttag 17662785  
caccacaatc gcagcagagg ctcaggatc cactgtaaacc tccgccttc aggttcaaac 17662725  
aattttctg ctcagccctc ccaagtagt ggaattacag gcgtgagcca ccacacccgg 17662665  
ccctgatttc ttaatatggc actcattata agattgtaaa agcccacctg tagaccaa 17662605  
tgggcacact ggctgctgc ttgtgacccctt tttccaggga aggacacagc tcccattagt 17662545  
ggctgaagta acacagttac aagaggcgga gttgggtttt gaactcagag ctccaggcgc 17662485  
cttacctta gggctcatcc ctttgagcaa aatgatgtttt cgaagagcat atcgttttaa 17662425  
ctgtggtttgc taatcaaggg gcctgattttt ggtgggaaat tcaactttaaac ttgttttaaa 17662365  
aggaaaacatt atgtcatcaa aatgggaaa ggcagttca cttgccataa ataggtcatg 17662305  
gtaaaaaaagt aaatgcaatg aaaacaacag tataattcaa tccaggctgg ttactattgc 17662245  
ctgcaggctg tgagactgtat tagtggtttt aacggaaagat gggcaaaaggc caggcagggt 17662185  
ttgcgagggc atgcccacact gaggcccttc taatatcatc aaggggtgg gggagccgg 17662125  
gwgccgtggc tggcctgtt aatcccaaggcctt cttctggagg ccaaggcttag gagaacactt 17662065  
gaggccggga gtttgagacc agtttggggca atacatgcag atcctgtctc tacaaaataa 17662005  
aaataaaaaaa cttagctggg ggttgtggta tgcacctata gtccttagcta cttgaaatgc 17661945  
tgagggcagga ggatcacctt agcccagaag ttcgagggtg cagttagctt tgggtgtcc 17661885  
actgcactcc agcctgggctt acagagcaag accttgcattt gcattttttt atttggtttat 17661825  
ttatattttt agacagggtc tcactcccat caccaggctt agagtgcattt ggcggaaatca 17661765  
aggctcatgt ccaccaatc ac cccctggct taagtgtatcc tcccacctca tgggtttgt 17661705  
attttgtaga gatgcggctt cactatgttgc tctaggctgg tcttgaactc ctgggtctaa 17661645  
gcaatccacc tgcctcaacc tcccaaaggat ctgggatttac aggctgtgaac caccacaccc 17661585  
ggccaagacc ctgtctctttt aatgtat aaaaaaaaaa aaggccgggg ggaagggtgg 17661525  
gggggaaatcc taaagaagag tttttctcac tctgagggtc acatcccttgc acccttgc 17661465  
caccctgtcc tgaagggtgt ctacccatc tgagctctcc ttgtactat cagtggctt 17661405  
ggaaacatgg ggatgtgttgc gtgtacatgtt cttatgtcc cttggccagg gggactggcc 17661345  
actgtccaca tggctggggg aggcttccccc ttctcagaag gcccacaaggc cagcagtgcc 17661285  
tacccatccc tggggcaggg gctgccacag gccaagtctt cagcctgtgg gagggtctgg 17661225  
ggctggccctt ggccttgagg tcaatggggat agcaggatgc tccctctgtt gtttcaagaa 17661165  
agaagctctt gtctgacagc atgcgtgggg aggcggggaa ggcctcggag gaggacactc 17661105  
tggtggtca cacgaggcgc tatcttaatg agctcaaggat acaggatgtc gggcctgggg 17661045  
ggctgcgggc ctggggcagg gggctgtgg ccaggagttt ccagaggcag gaggtgactc 17660985  
agcctgggg a gccaagtctt cacaggccac ccattcatgt cccttagtggt ggaggaacat 17660925  
gggagttgtt ggtccccaaag agaaggagag aggtcataaa aaggcagacc tcaatgggg 17660865  
ccaggccact ctgagggtgg tgcctccccc ttctccagggtt cgtatgaaag cttctataga 17660805  
attttaggt tctcattat gacttcaag ctgtgtctgtt ctcacacgc tccgagaccc 17660745  
cagccccctgtt ctcaccaacc tacatagctc ttctacttgc gtctatttt tttgtttgtt 17660685  
tggtcattttt tttttttttt tttttttttt aatggagtc tcaatgttc gcccaggctg 17660625  
gagtgcgtt atgtatctc ggtcttgc aacccctggcc ttccgggttc aagcaattat 17660565  
cctgtctcac cttctgttgc ggtcttgc aacccctggcc ttccgggttc aagcaattat 17660505  
aattttttttt gtatTTTttttagt agagatgtgg tttccgggtt tggccaaagc tggctctcgaa 17660445  
ctctctgttgc cttctgttgc ggtcttgc aacccctggcc ttccgggttc aagcaattat 17660385



ttactacctg tgatcacagg tgacatgtgg gaaggagtg agggagtgg tgatgtggc 17656605  
atctgggaa gggattcca agcagaagaa acagaagtg caaagatccc agggcagaac 17656545  
tatctgtcat gagttccagt atagtgtgg gagaaggaga cacagaccat agtccatgg 17656485  
agcacctgga gggaccctgg agagtctta gggagtgag ctcccttgg tctcaactc 17656425  
tctcttcctc tcctcgaggg getctctct ccttaaaaaa aaaattttt ttaattgtgg 17656365  
taaaattac ataacaaaat tcgcattaa ccactttaa ctgtacagg cagtggcett 17656305  
tagtccatc acaaagtgtc gcaaccatca tctctagttc caaacattt catcaactca 17656245  
aaaggaaacc ctgtgtcctt taaacactt ctccttattt atccccccaa gtcccttgg 17656185  
taatcactca cctgcattt ctctctatgg attgcttat cctggatatt tcataataat 17656125  
ggaatcatac aatatgtac ctttggc tggcttatct cactaagcac agcggtttca 17656065  
acattcatct gtgtgtgtt gtagatgtc tcaacttcc attccctttt acagcagaat 17656005  
gatattccat tgtaaaacac tacatttttt ttatccattt tagatgttat aggcccttgg 17655945  
gctattgtga gtatgttgc tggagatcg tgcatacgag tatttattt aatacctgtt 17655885  
ttcagttatt tgggtatac acctaggagt agaattactg ggtcacatgg taattctgtt 17655825  
taatttctg aagaaccatc aagggtatct ccacggggc tgcaccattt ccaccagtaa 17655765  
tgtaccagg tcccaattt ccatcttact ggatgtgaag tggtagtctca tggtttaat 17655705  
ttttttcccc ccaagtgtgg ttacacactg aggatctttt catgtgtga ttggcttattt 17655645  
ttgcatttac ctaatggcta gttttattca agtccttgc ccatttttaa aattggcttg 17655525  
gtatatgtca tttggagaaa gttctttata tattctggat attatTTTaa ttgtaaataa 17655465  
tctctcccat tctgtgggtt gttttttt tgatagtgtc ctttgatgca caaaaatttt 17655405  
agtttgcgtg aagtcattt ctttgcatttcc accgatgtg tttcttca agatTTTat 17655345  
ccattggccaa acccaaggc agtttcaactt ctttgatgttcc ttttgcatttcc agatTTTat 17655285  
tatatgtact tatatttagg ctttgcatttcc accgatgtg tttcttca agatTTTat 17655225  
caagtcccaac ttcatgttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17655165  
ctgttttcc cctgttgaat ttttgcatttcc accgatgtg tttcttca agatTTTat 17655105  
tatggattta ttcttagact ggtttgcatttcc accgatgtg tttcttca agatTTTat 17655045  
gggttgcatt ctttgcacag acctccgtct ctttgcatttcc accgatgtg tttcttca 17654985  
acctccgtct ctttgcatttcc accgatgtg tttcttca agatTTTat 17654925  
caggcgtgtg ctaccatgccc tggcttgcatttcc accgatgtg tttcttca agatTTTat 17654865  
catgttggtcgtt aggtggcctcc ttttgcatttcc accgatgtg tttcttca agatTTTat 17654805  
caaagtactg ggttacacagg atctatataatgtt ctttgcatttcc accgatgtg tttcttca 17654745  
atctatataatgtt caataccaca ggtttggatgg ctttgcatttcc accgatgtg tttcttca 17654685  
tttttggatgg ctttgcatttcc accgatgtg tttcttca agatTTTat 17654625  
tttttggatgg ctttgcatttcc accgatgtg tttcttca agatTTTat 17654565  
caactgcacc tccgccttcc ggtttcaggat gtttgcatttcc accgatgtg tttcttca 17654505  
gggattacag gggcccccgc tctcaccatg ttttgcatttcc accgatgtg tttcttca 17654445  
tctcaccatg ttttgcatttcc accgatgtg tttcttca agatTTTat 17654385  
ctcccaaagt gctgggattttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17654325  
taagatgtt ttttgcatttcc accgatgtg tttcttca agatTTTat 17654265  
ccatTTTTG caaaaaaaggc ttttgcatttcc accgatgtg tttcttca agatTTTat 17654205  
tggggagttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17654145  
tccgTTTTAT tatgtctttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17654085  
cttgcaccc ttttgcatttcc accgatgtg tttcttca agatTTTat 17654025  
tgaattttttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17653965  
tatagaataa caactgatttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17653905  
tttcttagca ttttgcatttcc accgatgtg tttcttca agatTTTat 17653845  
tgttaccagg ctggagtgca gtttgcatttcc accgatgtg tttcttca agatTTTat 17653785  
ttcaagcgat ttttgcatttcc accgatgtg tttcttca agatTTTat 17653725  
tgcccagcta attttgcatttcc accgatgtg tttcttca agatTTTat 17653665  
tctcgatctc ttgacccctgtt gatctggccca cttccgcctc tcaaagtgtc ggtattacag 17653605  
gcatgagcca ctggccctgg ctttgcatttcc accgatgtg tttcttca agatTTTat 17653545  
tgtgtgtgtg ttttgcatttcc accgatgtg tttcttca agatTTTat 17653485  
cgctgtgtaa gagaggttagc ttttgcatttcc accgatgtg tttcttca agatTTTat 17653425  
ctaattccctc tgatggAAC ttccagttact atgttaataa gcaagtagtgg agcaggcatc 17653365  
tttgctgtgt ttttgcatttcc accgatgtg tttcttca agatTTTat 17653305  
cagctgtgtgg gttaaattttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17653245  
ctagttttt gttttttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17653185  
catcaactga gagatcggtt ttttgcatttcc accgatgtg tttcttca agatTTTat 17653125  
aggaccacc taaagcaagc agtggggcgc ttttgcatttcc accgatgtg tttcttca 17653065  
agagcaggc ttttgcatttcc accgatgtg tttcttca agatTTTat 17653005  
gaaggtgaag gttttttt ttttgcatttcc accgatgtg tttcttca agatTTTat 17652945

cccaagtcatg	tcctgctctc	tgtggagtcc	cacagaggct	gacgaggat	gggggcctg	17652885
atacgctggct	acatgcaggc	catgccctt	ggcggttggt	ggcgtcagtc	tggggcagac	17652825
ctcccatgct	cacatagtgt	gctcattcac	ccagcaactgc	cttaggttgg	gctccctaga	17652765
atgggtggctc	ttaaacccca	gcaagtatct	gaaacactgg	agggcttgg	ccagcagatg	17652705
gctggggccc	tcccagagtt	tctgatccat	gttgtcttgg	gtagagactg	ggaatctgca	17652645
tttctaatac	attctcaagt	gttgtggatg	ctgtgtggct	gagaaccaca	tccctagaag	17652585
cagagtctga	gatgggtgcag	gcgatttcag	atgaaaccctg	caaggaggcac	aggcagtggg	17652525
gagccggcag	agtggcgcgc	tgagcacaga	tgtggattttgc	gaagtgtggc	ctcagcctga	17652465
ttccatggag	atctctgggg	cgtgaatgtc	accacagggt	tgccctggcc	agaagcatgt	17652405
ggccctggctg	ttagggcccc	ttgtcgtca	tggtctctct	gggatgtgc	aggtgaggtg	17652345
gctttctgtc	ggagaagggc	tctggtgac	cagccagaaaa	aggggatcaa	cggcatgcat	17652285
ggccagcacc	tactgtgtgc	cagggcatgge	ctcagcaactg	tctgcacage	agtgagcaga	17652225
cgcgtctgt	cctccctggag	ctggcatctt	tttgaggggag	atagatgta	atcggggacag	17652165
tctgttagct	cagggagagaga	agtgtatct	ggaagatga	agccaagtg	tgggtctccag	17652105
ggggcccccag	gtgggagtt	tttattttat	ttttttgaga	cagagttca	ctctgtcacc	17652045
caggctggag	tgcaagtgggt	cgatcttgc	tcactgcaac	ctccaccct	tgggttgaag	17651985
agattctctt	ccctcgccct	ctgagtagt	gggattacag	gcacctgcca	ccatgcccgg	17651925
ctaatttttg	tgtttttaat	ggacaccaga	tttcacccat	ttggccaggc	tggctgtgaa	17651865
ctctggacct	caagtaatcc	gcctacatca	gctctccaaa	gttctggat	tcacatgtta	17651805
agccaccaag	cctggctggg	tgtgggatt	tttagattaga	tgaggaggac	aggccctct	17651745
gactggttc	cacctctaag	tcctcatcca	aaggcttgg	ttatagatga	gacagaggca	17651685
cagagaagt	aattctaaat	tcacatagcc	agtggcagaa	cccagacttgc	gaccagtttgc	17651625
gggaacttct	gaggctgtcc	accccagtcc	tagcctcacc	cacagtccc	ttgcccaggg	17651565
cgagactatac	agggagcctg	acctgtgga	tctgggcagt	cccaccgtgg	catgtgcat	17651505
gtcccaagaga	aggatctgt	cagcagtgc	gcacccccc	cctgccccac	ccacagctcc	17651445
ctcggggct	atcccggaa	gtgttggta	gaaagtgtat	ctccagatgt	cacccgggttgg	17651385
tgcctgagc	tccctctacc	tgccaccttc	tctgaccaca	tagacccatgc	tctagcccaag	17651325
gccctctcc	ctctcctccc	ctcaccagg	gaccggccac	tagccccc	caccacttgc	17651265
gtttattttct	cacccggcc	actgtgggt	gttttcttcc	agagccgtgc	tgcctgtgg	17651205
aacccctgc	aatgtggaa	atgctcagac	ctgctctgt	cagtccagtc	gccactggcc	17651145
gcatgtggct	cttggaaat	ggagagtgt	actgaggaac	caaacttgc	ttttttaaat	17651085
tttgatgaat	ttacaatcac	tcgttaagtag	ccacccgtgg	ctggcagcc	ctggatttgg	17651025
ttgtgtctgg	ctagggtgtt	ggcaaccaca	tcactgcctt	gtgcagaaac	cactgtgcat	17650965
ccaggagaag	gcccaagtgc	cagccctc	ttcactgccc	gaagccctgt	gtcccgctga	17650905
ggggctctgc	tcgccaacgt	tggcacagca	aacacacata	cttctcttgc	tgggggctgg	17650845
tcctgtctgc	caagtccctgt	gcatgtctt	gggtggctgc	acccggccccc	tgcaccagg	17650785
caggccat	ctgtggagga	taccaaggaa	cctcttttgc	gttcccaatgt	gtgtcccat	17650725
ccactgcagt	tttgcagaaag	tttagtgcgt	gtgactttaa	aggccaaagg	ggcaggcaga	17650665
tcttctgaca	tctgggggaa	gcaaaatgttgc	aatggaaat	ttgctgcaga	atttctcaga	17650605
gccttttag	tgcttagat	tgctgc	atccaggagg	caggccgc	aaggccatgt	17650545
tcccaaacga	cttgcgggt	gaagcccttgc	tgaggagtg	tgtgcgagac	ccgtggctgt	17650485
ggagcacacg	agagaatgcc	tttctctgt	tttgtgttca	tgctggctc	tcggctgcat	17650425
tgtcttccag	tctgtgtccc	ctgctggctt	cccaggggagg	gagggaggt	gtgactccat	17650365
gtgtctcttc	agccgctgt	ttgttgc	attcgttcat	ggaaaaccat	gttccatgc	17650305
cagccacacg	cggggccctt	gcccggc	gggtgtggat	tgtgtaa	gaggagctga	17650245
tgacccatgg	cagggtaccc	ccttctct	ggctgttcc	gcaacatata	cacacgcaca	17650185
caacgcacacg	gacatacc	tgcacatc	tatacacaag	acacatatac	acacatatac	17650125
acactatgg	gtgtgtcc	cagctgtct	gtgtgtct	tcccagctt	tacactccca	17650065
ccccctccca	ggccctgt	tgcctccat	ttaccggcc	agggccctgg	cttgcggaa	17650005
tgggtccccg	tgggcaccc	tccttcc	ccatgtgt	gaccctgtc	actgcctt	17649945
ctaccatgg	ggggatgt	tgccagett	ccccgcctt	agccgcctt	gcccgcctgg	17649885
gctgggtggcc	atgggcattt	cccagcgt	ttggcaggct	gggtgccttgg	caccccccagg	17649825
actatgacag	aaggccccc	ttgtggcc	ggcctaagcc	atgaggcccc	tgctggggcc	17649765
tgactttaggg	tgtgtcc	cttttgc	ggccctgt	gctggctac	agcacccctt	17649705
ggccctctga	ggttcg	ccctct	tcacaccat	ccctggccac	cctctccct	17649645
cctgctgcct	gtgtctgt	attgaacat	ctcggtttt	tcccatcc	aaactccctt	17649585
tcctgggttgg	tgaacgc	ggccacactt	cccaat	tctcatgg	tgtctgc	17649525
ttgggtccctc	cctccac	ctcc	ccaccctt	tccacccgttgc	ctctgagca	17649465
ctgcacccat	ggtcttcc	catctc	ttcc	tgtcc	ttcc	17649405
gggttctctc	tctggcc	ggccctt	ccatgggg	aaggcccttgc	tcgtccccag	17649345
tcacccatcca	ctctcc	ccctgt	tcgttgc	tgccatgtt	actttaacc	17649285
cacccatggct	ccccaggggc	tgacttgg	ccatagagac	agaacctgt	gcccgccttgc	17649225

tacccctgtt caggttcacc tccaagtgcc attaccctca caggccccag acccgacacc 17649165  
tgggcctctt accccttgtc cctgcatgtc gcctgctaatt acctgtctt cttaccaccc 17649105  
cagacccttc ttatctcatg cttccctctt agggctgcta cttctctatt cctgttcccc 17649045  
taattgggtc tccttgctgc agctagtgca gcttggaca gcaccatcta tggttcccta 17648985  
ctggccctgac gacaatgtgt gagcctgtgc taggagacca ggcctgtgt gataagctca 17648925  
gcctgcctgt ttccagctgc acccaccctt tctatgtatcat ggactcaactt ctctgcccac 17648865  
agataccctt ttcccttgac ctctgcatct ggataactcc tattcaactct tcaccccttg 17648805  
caaatggcat cacccccaga aagecctt aataaaaaaaa acccagttt cctcttcata 17648745  
accacactca tcacactgca aataagtgtc tgcaagtgtc ctggcatgaa atggccct 17648685  
ccagtgccca cctggggcac ttagtataat atttacaaag tgagtggtc 17648625  
tgcctcgctg gggggggag cagggatgct tttttagcca ggagatggct tgggttgg 17648565  
gttcagctgg gcagccagtg ccatggatat ttacctgggt cacttggagg tcacaggga 17648505  
cactctgtcc tgatcttagt gcamatactt ttcaggtacc gtagacccccc ccagccttag 17648445  
cagctggaga tgagggcagt gcatccctt tgccaggaag gtccgattcc caatggacaa 17648385  
agaggcaatg cagtgcgagg ttagtataat tttttagggc ctgtccagct ctgtcaatg 17648325  
ctcattgtgt ggccttggga agatcctcgc tgccttaggca tcagtgcccc cttctgtaca 17648265  
gtgggtggtc tagactaatt tggtatccca aagcagtccct agacctgca tgcgtacttg 17648205  
gagccctctg cacccctgt tctgggcaca agagggcagc caagggccctt agaacgctga 17648145  
ggaaccctgg ccaactagct ttaagaaaatg cattgtgtaa actgtctttt acttgagccca 17648085  
gagcttgcca ggagcttggg agggttggg ctctgtctt cattttctacc aaaggaaatg 17648025  
tgcgttgcacca gggagtcat ccaaggggcac ctggaaacttgc tccctcaaggc atttccggg 17647965  
gaaccaattt ctcacgggtt ttcggcagttt ggggaaagcgg aggccaaacag cccctgtt 17647905  
tttccgcagt ggtcccttgc ccaacttcc ttgtgcagag gaagggtgtg aggcccttc ggacccagac aggagggacc 17647845  
ataatggtag gtgggggtggg ttctcacctc ctttgcctgtt gaatggccctc ctcccactta gtagttgaac agaatctaa 17647785  
atattccctca aggtcggca ggacatcaga attctttctt ttttttctt cttttttttt tgagttagt 17647725  
ttgcttagtat ataataatacc aaggcacacc ccctcaccag ctcgttctgt cgccccaggct tttccgcactt ggtgtatct ttgtgtactg 17647665  
tttcccttaggtt caagcgattt cccacccatg cttgtctaat cttttttttt tgagttagt 17647605  
ggggcatgac gggggcatgac tttttttttt tgagttagt 17647545  
tttccaccca ccaaaagcgga tttttttttt tgagttagt 17647485  
ttttttttttt tttttttttt tgagttagt 17647425  
atattccctca aggtcggca ggacatcaga attctttctt tttttttttt tgagttagt 17647365  
tttcccttaggtt caagcgattt cccacccatg cttgtctaat cttttttttt tgagttagt 17647305  
ggggcatgac gggggcatgac tttttttttt tgagttagt 17647245  
tttccaccca ccaaaagcgga tttttttttt tgagttagt 17647185  
tttccaccca ccaaaagcgga tttttttttt tgagttagt 17647125  
tttccaccca ccaaaagcgga tttttttttt tgagttagt 17647065  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17647005  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646945  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646885  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646825  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646765  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646705  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646645  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646585  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646525  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646465  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646405  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646345  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646285  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646225  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646165  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646105  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17646045  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645985  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645925  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645865  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645805  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645745  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645685  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645625  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645565  
tttccaccca ccaaaatagg tttccatgg tttttttttt tgagttagt 17645505

gaggcacacgc acttgacaat ttacaaagct cttttcacc aggtctttt tttcttttc	17645445
gagaacctgtt ttcactcttg ttgcccaggc tggagtgc当地 tggcgcgatc tcggctcacc	17645385
gcaacccctcg cctccccaggc tcaaacaatt ctctgc当地 agcctcctga gtagctgaga	17645325
ttacaggcat gcaccaccaat gcccggctaa ttttgtat当地 ttagtagaga cagggttct	17645265
ccatgttggc caggctgggt cttgaactcc cgacccagg tgatacgc当地 acctcgctcg	17645205
gcctccaaa gtgctaagat tacagacatc agcaccacag cccggccatc acccagactc	17645145
ttatttgagc tgggcataat tgtcaggct gtc当地 ctactgtc tgagggaaa gccatggaaa	17645085
gatgctact ggatcgatc gagccctaaa gcagggtccc ccagccctt gctctgact	17645025
ctgcagggg gatccacct tggccactg cacagttgg gggagcccca ctctgcaggg	17644965
gctgggtctc ttccatcttg gtattaccag gtgc当地 tagctgg catagtaatg	17644905
atgttatggt actctgtgc acaaaccgg gagtgtatc tgccctgc当地 gtctacagca	17644845
gggttccggag gaggccctgg atggccctcc ccatggcagg tgtaactgcc tggttagggt	17644785
taagagctg gatcctgatc caccctggg ttgatcctgg ttctgc当地 acctggctgt	17644725
gtgaccctgg gcaagttgtc gacccctctc gtgggtc当地 agtctatct gtaaaatggg	17644665
gatgggtatc ctaatgc当地 ccctcggtc ggagggagtc tttagc当地 agtctggc	17644605
agtcaagggt tcaactgtggc tgc当地 tctactgtc tcattaggag ccaacagtg cctctgggg	17644545
ggtggggagag gcaagttcct ggtatccatc gggccgtc当地 gacactgtc gacggagcag	17644485
ttgttgggct caatttccatc gggccctc当地 aattcaggcc atcccagggg ctgcagggg	17644425
gggggtatct atgggc当地 ctttgc当地 gctgtgtc当地 agggttggg ggtgatggat	17644365
cccggtctc aggccctcc tcgtggctgt aggcaactatc gaccaggcaga ggggtccctt	17644305
cctgaccacc cgcttggcc actggcagaa tccgtgtggc ccccatatcca ccactccctc	17644245
ctggagttggg gagccacatc gagccaggcc cagcttggg gggacaagga gcagcttct	17644185
gcttctggaa ttagtgc当地 tctgttgc当地 aggggtgtga gtggcactga ggacttgc当地	17644125
gggacaccctt gaagatgtgg ctgc当地 tctggggat ggtgacatgc cccagcactc	17644065
agcttagttt gccaacccag agtccgaggc acagggttct gagactgtag cagggaggat	17644005
gctggggagat gtgaagggat ggaggagctc ctggactgag cactgtctc当地 tggctctgag	17643945
cagaccgc当地 ctctgc当地 ccgc当地 tggcttccac cactgtctc当地 ggc当地 ctgg	17643885
cgggggcttc tggcctatc cggacatcagc gtc当地 cc当地 tggctctc当地 aagggtgtc当地 tatgagcaag	17643825
tggggctctc cttccaaagat ccctctggat atccctccatc tagctccaaa ttaactgttc	17643765
tcaccctgaa ttatagacaa gggccatc当地 ctggaggcagg gaggggctt gtttgggtt	17643705
ctcagccagg ctggactgtc atccagatct gacacttgc cctcttccat gttgctttaga	17643645
aggggtgc当地 gtgggtggaa ggagttattc cagccctccca cagagccagg ggactagaga	17643585
gggtcaggat ctgctgtata gccc当地 atttgc当地 aagttgtagg aagaaggc当地 tggctggcaa	17643525
agggagtagg gagtggaaag aatgatgtg ctgatagcac ctggcaggcc tgc当地 ctcc	17643465
aaccgc当地 gtgctccagg acttactccc tgaatcctcg cagacagaca gggccc当地	17643405
gaggtgaggg catgcaaata gcaggggcag aatttgc当地 ggc当地 ctggggcc	17643345
ccacaactcc ctggccactc tggcctggc cttgtgtgg gcatcaggaa ctgactgacc	17643285
tggcctatc tggcctgtc ctcatggggc acatagactatc atggggggaa gcaaggccatt	17643225
aggagaaggg gaaaggcagc gagacccctc tggggaggag ggaatgaagg cttctggaa	17643165
gagggggcat ttaggacttg gccttgc当地 ataaaggcaga ggttgggac tgaagtc当地	17643105
gggctgtggg gattctctcc ttaaccctca cacatttctc agggatctg gaaaaatcca	17643045
gggctctggat gaccacttca cctcttgc当地 tatgaccctt cagggcacag gacatgc当地	17642985
ctcccccagg gageccccc tgaccaccc ctgcatgc当地 acatggagcc ccacagctgg	17642925
agctgc当地 ctctccctgg caagtgc当地 cttgtgtggg tggcctgatt accccacaagc	17642865
attagggccc cctccccc当地 cctcgccaggc cagctggggag ttgtgttagg gctgggtct	17642805
ctgtccccc当地 cagatcctca tggc当地 acatccctc当地 gcatgttctg tttagc当地	17642745
tggagggcat ctccagggtt accatcattt atttgtatc ccatcagggt agtccctgc	17642685
agggctggc当地 ctcttagggg acctggccacc cccaggcttca gaatctccc gggcaggag	17642625
agtctccctc当地 ctatgtccc caggctctc acggcttctg tcttctgtct ctcggctac	17642565
aaatcgaggg tctgtctttg tcaactgtc当地 caggacagcg ggtcctctc当地 attgtccc当地	17642505
agggtctcc local ctccctctc当地 ctgactgc当地 ccacatgagg ctcttc当地 agccactct	17642445
gatggactg ctctcgatc当地 cagagctctg ctgtgggtcc ccattgttta tgaataattt	17642385
ggggcactgc cccctgccc当地 gagctgtc当地 gcactggcc当地 cctgccc当地 aggccgatgc	17642325
ccacacacat gcttggctc gggcacctgg ggtcaccatt taagaactcg ggc当地 ctgg	17642265
agtaaagtgt caaaggcagag ggtaacctcc tccctaggac ccctaatgag gccactgc当地	17642205
ctggc当地 agggaggggg cccaggctggc tccggaaaggc gcaaggcc当地 acccactactg	17642145
ctggc当地 gtgtctgtt当地 gggccctgccc ttgggttctt ggc当地 cc当地 actctgagc	17642085
caggctcacat gtggacagtc ct当地 tttacaggc当地 tgctttctc当地 atccctgatc ccaaccagtc	17642025
ccaccacaga ctttgc当地 gaggggggg tggcaggc当地 ggatttctc当地 ctctgatagg gaaacctaaaga	17641965
gcactgggt tggc当地 aggggggggg gatggc当地 ggatgttta aggttgaatc	17641905
ccagctctc当地 cc当地 taaacag tcatgagacc tggcctcccc当地 gagggcaggc cgtgctgccc	17641845
tggcaaatttgg gagtttctt当地 gaggggggggg tgggtggc当地 agccccagcc ttgc当地 ctgg	17641785

cacctacccg agagcggcta ctgtgaccts cccacagggc aatgggcattg agcgagactt 17641725  
catggacgac aagcgtgtgt acatcatggaa tgtctacaac cgccacatct acccaggggaa 17641665  
ccgcttgcc aagcgtaagg tgcgtccccct accctcatct tgggtgtgtc cttgtggatg 17641605  
aggctctctc ctgagtgtct cctgtctgct aggccctgca gaagccactg cagtggttca 17641545  
tagcatccct gtgaggtgtat cctttccatt ttacagatga ggaaaccggag acctggagaa 17641485  
gtcactcgac ccacccaaga tcacataacc cttacaataa acatgcattt gtctggcaaa 17641425  
aaacaggaaa gaatggaaa aaaaaaagaa aaataggata aatttggaaaa taagaaataa 17641365  
gaaataattt cacataggct gggcgccgtg gtcacgcct gtaatcccag cacttggga 17641305  
ggctgaagtg ggcggatcac ctgaggctgg gagtttgaga ccagcctgac caacatggag 17641245  
aaaccccatc tctactaaaa atacaaaatt agctggatgt ggtggcgtat gcctgtatc 17641185  
ccagctactc gggaggctga gacaggagaa ttgttgaac ctggggaggcg gaggttccgg 17641125  
taagccgaga tcgcgccatt gcactccgc ttgggcaaca agagcgaac tccatctcg 17641065  
aagaaagaaa gaaattcatg tataatcggtt aaaatggaaa tgcatggaaa tcatcaatca 17641005  
aaaggcagag actctcagat gagattttaa aacagggtctg ccaccttgc aggtagggaa 17640945  
cattttgca ccagtcacga tgagtctgggt gtggataagt cagcagctag tatggcccaa 17640885  
ggaaccaatt tctgaacaga acctcacatg tgctgagcct gggcttaagg gcagggcagg 17640825  
gtgtccatgt gtgtaggcaa gacccagagg aggcaagtggaa atctgacatt gccgacacag 17640765  
atctccacac ccccgaggca gtgtctcgc ttcaatggccc cttctctctt ttgatcccc 17640705  
cttttgca ctcttgcgtc ttcttgcacc tttagttttgg gtggaatggag gtcgagcgt 17640645  
gctgaatctg acagaccagt ttccagctt gcctgggtgc cacagtcttgc ttctgagcct 17640585  
cagttccct tctctataaa ttgaggccat ccatgtctct ctcccagagg ccattcaggcg 17640525  
gaagggtggag ctggagttgg gcacagagga tgatgagttac ctggataagg tggagaggaa 17640465  
catcaagaaa tccctccagg agcacctgcc cgacgtgggtg gtataacaatg caggcacccg 17640405  
catcctcgag ggggaccgc ttgggggct gtccatcagc ccagcgttac gcctgaccc 17640345  
ttggggccac gggagggtct gctctatggaa ctcagcagca gcaggaaagg tggggccct 17640285  
catgtcagggg aggagatggaa ctgaagaac agcagttgg agcagggtta gcccctgcgc 17640225  
aggacttcct gacaccatgg gggtctggcc tgcctgagtc accctccctt tccccctaaaca 17640165  
gggcatcggt aagcgggatg agctgggtt ccggatggtc cgtggccgc gggtgcccat 17640105  
ccttatggtg acctcaggcg ggtaccagaa ggcacagcc cgcatcatttgc tgcactccat 17640045  
acttaatctg tttggcctgg ggctcattgg gcctgagtc cccagcgtct ccgcacagaa 17639985  
ctcagacaca ccgctgcttc cccctgcagt gccc

- 1 -

SEQUENCE LISTING

<110> Novartis AG

<120> Histone Deacetylase-Related Gene and Protein

<130> Case 4-32094A

<160> 4

<170> PatentIn version 3.0

<210> 1

<211> 347

<212> PRT

<213> Homo sapiens

<400> 1

Met Leu His Thr Thr Gln Leu Tyr Gln His Val Pro Glu Thr Pro Trp  
1 5 10 15

Pro Ile Val Tyr Ser Pro Arg Tyr Asn Ile Thr Phe Met Gly Leu Glu  
20 25 30

Lys Leu His Pro Phe Asp Ala Gly Lys Trp Gly Lys Val Ile Asn Phe  
35 40 45

Leu Lys Glu Glu Lys Leu Leu Ser Asp Ser Met Leu Val Glu Ala Arg  
50 55 60

Glu Ala Ser Glu Glu Asp Leu Leu Val Val His Thr Arg Arg Tyr Leu  
65 70 75 80

Asn Glu Leu Lys Trp Ser Phe Ala Val Ala Thr Ile Thr Glu Ile Pro  
85 90 95

Pro Val Ile Phe Leu Pro Asn Phe Leu Val Gln Arg Lys Val Leu Arg  
100 105 110

Pro Leu Arg Thr Gln Thr Gly Gly Thr Ile Met Ala Gly Lys Leu Ala  
115 120 125

- 2 -

Val Glu Arg Gly Trp Ala Ile Asn Val Gly Gly Gly Phe His His Cys  
130 135 140

Ser Ser Asp Arg Gly Gly Gly Phe Cys Ala Tyr Ala Asp Ile Thr Leu  
145 150 155 160

Ala Ile Lys Phe Leu Phe Glu Arg Val Glu Gly Ile Ser Arg Ala Thr  
165 170 175

Ile Ile Asp Leu Asp Ala His Gln Gly Asn Gly His Glu Arg Asp Phe  
180 185 190

Met Asp Asp Lys Arg Val Tyr Ile Met Asp Val Tyr Asn Arg His Ile  
195 200 205

Tyr Pro Gly Asp Arg Phe Ala Lys Gln Ala Ile Arg Arg Lys Val Glu  
210 215 220

Leu Glu Trp Gly Thr Glu Asp Asp Glu Tyr Leu Asp Lys Val Glu Arg  
225 230 235 240

Asn Ile Lys Lys Ser Leu Gln Glu His Leu Pro Asp Val Val Val Tyr  
245 250 255

Asn Ala Gly Thr Asp Ile Leu Glu Gly Asp Arg Leu Gly Leu Ser  
260 265 270

Ile Ser Pro Ala Gly Ile Val Lys Arg Asp Glu Leu Val Phe Arg Met  
275 280 285

Val Arg Gly Arg Arg Val Pro Ile Leu Met Val Thr Ser Gly Gly Tyr  
290 295 300

Gln Lys Arg Thr Ala Arg Ile Ile Ala Asp Ser Ile Leu Asn Leu Phe  
305 310 315 320

Gly Leu Gly Leu Ile Gly Pro Glu Ser Pro Ser Val Ser Ala Gln Asn  
325 330 335

Ser Asp Thr Pro Leu Leu Pro Pro Ala Val Pro  
340 345

<210> 2  
<211> 1755  
<212> DNA  
<213> Homo sapiens

<400> 2  
agctttggga gggccggccc cgggatgcta cacacaaccc agctgtacca gcacgtgc 60  
gagacaccct ggccaatcgt gtactcgccg cgctacaaca tcacacctcat gggcctggag 120  
aagctgcata cctttgatgc cggaaaatgg ggcaaagtga tcaatttcct aaaagaagag 180  
aagcttctgt ctgacagcat gctgggtggag gcgcgggagg cctcggagga ggacctgctg 240  
gtgggtgcaca cgaggcgcta tcttaattag ctcagaatgggt cctttgctgt tgctaccatc 300  
acagaaatcc cccccgttat cttccccc aacttccttg tgcagaggaa ggtgctgagg 360  
cccccttcgga cccagacagg aggaaccata atggcgggga agctggctgt ggagcgaggc 420  
tgggccatca acgtgggggg tggcttccac cactgctcca gcgaccgtgg cgggggcttc 480  
tgtgcctatg cggacatcac gctcgccatc aagtttctgt ttgagcgtgt ggagggcatc 540  
tccaggggcta ccatcattga tctttagtgc catcaggga atgggcatga gcgagacttc 600  
atggacgaca agcgtgtgt aatcatggat gtctacaacc gccacatcta cccaggggac 660  
cgctttggca agcaggccat caggcgaaag gtggagctgg agtggggcac agaggatgt 720  
gagttacctgg ataagggtggaa gaggaacatc aagaaatccc tccaggagca cctgcccac 780  
gtgggtgtat acaatgcagg caccgacatc ctcgaggggg accgccttgg gggctgtcc 840  
atcagccccag cgggcacatgt gaagcgggat gagctgggtgt tccggatgtt ccgtggccgc 900  
cgggtgccccca tccttatggt gacctcaggc gggtaaccaga agcgcacagc cccatcatt 960  
gctgactcca tacttaatct gtttggcctg gggctcattt ggcctgagtc acccagcgtc 1020

- 4 -

tccgcacaga actcagacac accgctgttt cccccctgcag tgccctgacc cttgtgtccc	1080
tgcctgtcac gtggccctgc ctatccgccc ctttagtgctt tttgtttctt aacctcatgg	1140
ggtgtgtggag gcagccttca gtgagcatgg agggggcaggg ccatccctgg ctggggcctg	1200
gagctggccc ttccctact tttccctgct ggaagccaga agggcttgag gcctctatgg	1260
gtgggggcag aaggcagagc ctgtgtccca gggggaccca cacgaagtca ccagccata	1320
ggtccaggaa ggcaggcagt taactgagaa ttggagagga caggctaggt cccaggcaca	1380
gcgagggccc tgggcttggg gtgttctgg tttgagaacg gcagacccag gtcggagtga	1440
ggaagcttcc acctccatcc tgactaggcc tgcatcctaa ctgggcctcc ctccctcccc	1500
ttggcatgg gatttgctgc cctcttgcc ccagagctga agagctatacg gcactggtgt	1560
ggatggccca ggaggtgctg gagctaggc tccaggtggg cctggttccc aggcagcagg	1620
tgggaacctt gggcctggat gtgagggcg gtcaggaagg ggtacaggtg ggttccctca	1680
tctggagttc cccctcaata aagcaaggc tggacctgca aaaaaaaaaa aaaaaaaaaa	1740
aaaaaaaaaa aaaaa	1755

<210> 3  
<211> 1044  
<212> DNA  
<213> Homo sapiens

<400> 3	
atgctacaca caacccagct gtaccagcat gtgccagaga caccctggcc aatcggtac	60
tgcgcgcgt acaacatcac cttcatggc ctggagaagc tgcatccctt tgatgccgga	120
aaatgggc aagtgtatcaa tttcctaaaa gaagagaagc ttctgtctga cagcatgctg	180
gtggaggcgc gggaggcctc ggaggaggac ctgctggtgg tgcacacgag gcgctatctt	240

- 5 -

aatgagctca agtggtcctt tgctgttgct accatcacag aaatcccccc cgttatcttc	300
ctccccaaact tccttgtaca gaggaagggtg ctgaggcccc ttcggaccca gacaggagga	360
accataatgg cgggaaagct ggctgtggag cgaggctggg ccatcaacgt ggggggtggc	420
ttccaccact gctccagcga ccgtggcggg ggcttctgtg cctatgcgga catcacgctc	480
gccatcaagt ttctgtttga gcgtgtggag ggcatttcca gggctaccat catttatctt	540
gatgccatc agggcaatgg gcatgagcga gacttcatgg acgacaagcg tgtgtacatc	600
atggatgtct acaaccgcca catctaccca ggggaccgct ttgccaagca ggcattcagg	660
cggaaaggtagg agctggagtg gggcacagag gatgatgagt acctggataa ggtggagagg	720
aacatcaaga aatccctcca ggagcacctg cccgacgtgg tggtatacaa tgcaggcacc	780
gacatcctcg agggggaccg cttgggggg ctgtccatca gcccagcggtt catcgtaag	840
cgggatgagc tgggtttccg gatggtccgt gggccggg tggccatct tatggtgacc	900
tcaggcggtt accagaagcg cacagccgc atcattgctg actccataact taatctgttt	960
ggcctggggc tcattgggcc tgagtacccc agcgtctccg cacagaactc agacacaccg	1020
ctgcttcccc ctgcagtgcc ctga	1044

<210> 4

<211> 23434

<212> DNA

<213> Homo sapiens

<400> 4

ctacacacaa cccagctgta ccagcatgtg ccagagacac gctggccaaat cgtgtactcg	60
ccgcgcatac acatcacctt catgggcctg gagaagctgc atcccttga tgccggaaaa	120
tggggcaaag tgatcaattt cctaaaaggt atggaaggtc ccccttggac tctcatctgc	180

ttcctccaac ccacctgtcc tctccgtcct catccccaaac ataaggctca ggctcttc 240  
catcttcagt ttcagccctc gnatggcctt ccacccatgc ttccgcccaa aatgatttt 300  
ccaacacaga ctcctaataca cgatatgatg tccctgactc agactctccc tggctcccc 360  
tcctgtggc ctaagtcctg cctctgccc agaggcctag tggaaaggta gctgattact 420  
gatggcaca gggaaaggta agcttggagg agtccatttc ctaaggttca gagagtcagg 480  
agtagagca cctccacccgc acctctttt attacagatg gggaaattt 540  
tgattaggaa acatgtgcac ccaattccag tccagtcctc acagcagccc tcgggttagg 600  
caccacaatc gcagcagagg ctcaggagct cactgttaacc tccgccttc agttcaaac 660  
aattttctg ctcagccctc ccaagtagct ggaattacag gcgtgagcca ccacacccgg 720  
ccctgatttc ttaatatggc actcattata agattgtaaa agcccacctg tagaccgaac 780  
tggcacact ggctgcctgc ttgtgacctc tttccagggaggacacacagc tcccattagt 840  
ggctgaagta acacagttac aagaggcgga gttgggttttgaactcagag ctccaggcgc 900  
cctaccttta gggctcatcc ctttgagcaa aatgatgctt cgaagagcat atcgtttaa 960  
ctgtggttttaatcaaggg gcctgattta ggtggaaat tcacttaaac ttgtttaaa 1020  
aggaaacatt atgtcatcaa aatggaaaaa ggcagttca cttgccataa ataggtcatg 1080  
gtaaaaaaaaatgaaatgaaaacaacag tataattcaa tccaggctgg ttactattgc 1140  
ctgcaggctg tgagactgat tagtggttttgaacggaagat gagcaaagca caggcagggt 1200  
ttgcgaggcc atgccacact gaggcctcctg taatatcatc agaagggtggaa gggaggccgg 1260  
gcgcaagtggc tcgtgcctgt aatcccagca ctctggagg ccaaggctag gagaacactt 1320  
gaggccggga gtttgagacc agcttggca acatagcaag atcctgtctc tacaaaataa 1380

- 7 -

aaataaaaaa cttagctggg ggtggggta tgcacctata gtcctagcta cttgaaatgc 1440  
tgaggcagga ggatcaacttgc agcccagaag ttgcagggtg cagttagctt tggttgtgcc 1500  
actgcactcc agcctgggtc acagagcaag accttgcattt gcattttttt atttttttat 1560  
ttattttatttgc agacagggtc tcactccat cacccaggct agagtgcagt ggcggaaatca 1620  
aggctcatttgc ccaccaaac ctccctggct taagtgtacc tccccacca tgggtttgtat 1680  
attttgtaga gatgcgggtct cactatgttgc tctaggctgg tcttgaactc ctggggctcaa 1740  
gcaatccacc tgcctcaacc tcccaaagtgc ctgggattac aggctgtac caccacaccc 1800  
ggccaagacc ctgtctctttt aatgaatta aaaaaaaaaa aaggcgcccc ggaagggtgg 1860  
gggggaattt ctaagaagag tttttctcac tctgggggtc aacatccctg acccttgc 1920  
cacctgctcc tgaagggttgc ctgcacacc tgcgtctcc ttgtgactat cagtggttgc 1980  
ggaaacatgg ggattgtgt gtgtacgttgc ttcatgttgc cctggccaga gggactggcc 2040  
actgtccaca gtggctgggg aggctacccc ttctcagaag gcccacaagc cagcagtggcc 2100  
tacctacccc tggggcaggg gctgccacag gccaagtctg cagcctgtgg gagggtctgg 2160  
ggctggccct ggccttgagg tcagtggga agcaggatgc tccctctgtg gtttcagaag 2220  
agaagcttct gtctgacagc atgctgggtgg aggccgggaa ggcctcgagg gaggacctgc 2280  
tgggtgtca cacgaggcgc tatcttaatgc agctcaaggt acaggatgc gggcctgggg 2340  
ggctgcgggc ctggggcagg gggctgtgg ccaggagggtt ccagaggcag gaggtgactc 2400  
agcctggggaa agccaaatctt cacaggccac ccattcatgt cccttagtgg ggaggaacat 2460  
gggaggtctgt ggtccccaaag agaaggagag aggtcataaa aaggcagacc tcagtttggg 2520  
ccaggccact ctgagggtgg tgcctcccc ttctccagggtt cgtatgaaag ctttcataaga 2580  
attttaggct tctacattat gactttcaag ctgtgtctgc tcgacacgccc tccgagaccc 2640

cagccccctgt cctccaacca tacatacgctc tttcaacttg gtctatttg tttgtttgtt 2700  
tgttcatttt tttgttgtt tttgttcttg aaatggagtc tcactctgtc gcccaggctg 2760  
gagtgcagtg atgtgatctc ggctcattgc aacctccgcc ttccgggttc aagcaattat 2820  
cctgtctcac cctcctgagt gagtagctag gattacaggc gcgtgccacc atgcctggct 2880  
aattttttt gtattttagt agagatgtgg tttcgccgtg ttggccaagc tggtctcgaa 2940  
ctcctgaact caggtcatct gcccacctcg gcctccaaa gtgctgggt tacaggcgtg 3000  
agccaccgca cccaccctat ttttatatt gggctgaagt ttaagactct ggtctaagta 3060  
cttctgctga agttttgttg aaaattgttg gtctaaaaac taatttggaa ccctcagggc 3120  
tcagcagaga agagaaaacaa gtgggagggc cggtggtaga gtctgaggtg aactcctgcc 3180  
cttcccaag gggcggtcc tcagctccac tgtggcccg gcatggccag agcacctgg 3240  
cttcaaagag aagccagggaa tccagattat taagtgacat ttcctgattt tttttttag 3300  
actgagtctc gctcttggc agcaggctga agtgcagtg cacgatctca gctcactgta 3360  
acctccgcct cccgggttca aacaattttt ctgcctcagc ctccgaagta gctggatta 3420  
taggggttag ccaccacacc cggccctgat ttcttaatgt ggcactcatt ataagattgt 3480  
aaaagccac ctgttagacca aactgggcac actggctgcc tgcttgac ctcttccag 3540  
agaaggacac agtccttatt agtggctgaa gttctgaggg ctgaggcatt cagttcagtg 3600  
ctctttgttag gaacagaggg gaggtgggg cgggggcttg cattgaaatc tggtaactgcc 3660  
agcctgcctt ggtgggggtg gggtcaggga tgcctcaggt tatctgcccc aagagtgtgg 3720  
gagccctgac ccccaggctc cctggctgag ctcacccctag actcagagcc acagtggatg 3780  
cctgaggcca gcaggccct ctgctccaca ggtggaaaag cctaggtccca gaaagaggct 3840

- 9 -

tggtcaagg tcacctggga agttggccgg gccttggga gacccctggc aggtcatcca 3900  
gtccagtctt ctaggcccc agtcagggct gctgcctccc tgctcccaa ccgcagcctg 3960  
aggtgtgaga attctagata gggccacgac agtgtgagca catgaaagat taccaggaag 4020  
aggtgaaac ctggctcctg ggagagagag ggggtgtgagg cttggcagg aagccagtg 4080  
cttggctgcc ctggttcct ggggcccagg catgcgtggt cacagtccac agcctaggc 4140  
tggccagga ggacatgcct gccagagtcc cgagggtgag gggaaaggaag ggacaggagg 4200  
cgctcagctg gggcagggag aaaccaaaac agaatggtgt gattgaacca ggctgggggt 4260  
ggggggccta gttccaggg ccccacccat ttgaggggcc ttcagggaa ctgtgttggg 4320  
caggctgcat gcctggcctt ggtcccaa aagcctgaaa gcagcttaact atgtgatata 4380  
taataataca aaatagctgg gtgttagtggc atgcacttgt agtcctagct acttgggagc 4440  
ctgaggcagg agacctttag cccaggagtt tgaagctgta gtgagctatg attgcaccac 4500  
tgcactccag ctggtatgac agagttagac tgtctcttaa aaaaataat aaaagtatta 4560  
acaggttagag tcccaagtag aaaaactgagg ttgagggttag gaggagaatt caggtatgtc 4620  
cactgaaaaa gttAACCAAG atggtgatcc agctgcatat ttggcttggc gctccctggc 4680  
agtcagaaca aaaggagaaa catgatggtt tctacggcac ctattaagat gaagaagtag 4740  
gccgggtgca gtgactcatg cctgtaatcc cagcactttg ggagaacgag gcggggcggat 4800  
cacttgaggt cgggagtttgg agatcagccct ggccaacatg gagaaacccct gtctctacta 4860  
aaactacaaa attagccagg catggtagtg catgcctgta atccagcta cctgggaggc 4920  
tgaggcagga aaatcacttg aacctggag gtagaggttg cagttagccg agattgcgcc 4980  
attgcactcc agcttggca ataagagtga aactccatct caaaaaaaaaa aaaaaaaaaa 5040  
aaagaaaaaaag atgaagaagt agtcagtcataaacacatc tgtattgaat gccaactgta 5100

- 10 -

cagagagaat aagacagcag ggctctctgc caccatggat ttgcatttga gtcgtggaag 5160  
ataaaaatta aggaagcaac cacccaaagag catttttagag agcaccaagg gctatgaaga 5220  
aagtgaaaaa tagagggtaa ttggatggtc agggagggcc tcacagagga ggtgatgtt 5280  
gagttgagac taaacaaagg agcaggtgat actcatgtag aggtgtttt tttttttt 5340  
ttttttttt gagaaggaat ctcgctttgt tgcccaggct cgagtagt ggtgcgatct 5400  
cagctcacag caacctctgc ctcttggttc aagcgattct cctgcctcag cctcccaagt 5460  
agctaggatt acaggcacct gccaccatgc ccggctaatt tttgtattt tagtagagac 5520  
ggagtttca ccatgttggc caggctggc tcgaactcct gacctcaagc aatccatctg 5580  
cctcggcctc ccaaagtgct gggattacag gcatgagcca ctgctcctgg ccctcatgta 5640  
tagcttgaa ggaagaatgt ttcagaatcc caggcctgga gggtggaggg gacttgatct 5700  
tccaaagggg agaagaatgc ttgggaggcc ggatggaagg gaataaaaca ttgtggctcg 5760  
tacacggtgc agttagggag gccagagccc cagggcacac aaggcttgc aggccgtggg 5820  
aggagtgtat atgttgttcc agggacctt gacagtcacg agggggttt cagcaggagg 5880  
gtgatatggt gtgacatgcc cttgctgccc aggtggacc caagcccggt tcagacatca 5940  
tctggcacct aaggctgcag ctcaggaaca tctcccacct ccctgcagat gtctgcaatg 6000  
tttctttct ctttcctctg ctgtggcgc ccagagagtg ccctagagag tccttcaggt 6060  
ttctcaggct gctttccct ggtcattctg tgtgtgctgt gtaacatcca ccgtctcccc 6120  
tgccctcatcc cattctaccc ccaaccctg cctggggctc atgcctgact ctgcactgg 6180  
gtggccttg atacttaata aacagggcac tgaaggagaa gcaggagctg gacgtttgca 6240  
agatgtcaat tcagggaaac ccatgtttat caagctcctg ctgtgtgcaa ggtccagggt 6300

- 11 -

tggccctct gaggtagct gttgagctcc ccagtcccc agcaactggc tcttcctt 6360  
ggttgcattct cgggtgacag tttgcacatgg agtgcgttt agtgcgtggc agcatctgac 6420  
atcctgcccc tgtgcactct gcactggaca gtgctcagaa cacgtggatc cagcaagtgc 6480  
tcagagggca ccactctgtg atctaggatc tgcaaggatg ggatggagca aaagaccaca 6540  
tcccttcct gctggagctg gcatttaggt gggagagtca gacaataaat gtaataatta 6600  
agtaatgaga taatatgtta gatggtgctg agtgcgtga agaaaggaag ggacacgaga 6660  
aaaggggtgg ggagagctgg tgagaggatg gcagtttaa atcaggatc aggaaaggc 6720  
ttactacctg tgatcacagg tgacatgtgg gaagggagtg agggagtggg tgatgtggc 6780  
atctgggaa gggcattcca agcagaagaa acagcaagtg caaagatccc agggcagaac 6840  
tatctgtcat gagttccagt atagtgtgga gagaaggaga cacagaccat agctccatgg 6900  
agcacctgga gggaccctgg agagtctcta ggggagtgag ctcccttgg tctccaactc 6960  
tctcttcct tccctgaggg gctcctctct cctttaaaaa aaaattttt ttaattgtgg 7020  
taaaatttac ataacaaaat tcgccattaa ccactttaaa ctgtacagtt cagtgccctt 7080  
tagtccattc acaaagtgc gcaaccatca tctctagttc caaacattt catcaactca 7140  
aaaggaaacc ctgtgtcctt taaacacttg ctccccattt atcccccaa gtcccttgg 7200  
taatcactca cctgcattct ctctctatgg atttgcctat cctggatatt tcatataaat 7260  
ggaatcatac aatatgtgac cttttgtgtc tggcttatct cactaagcac agcgtttca 7320  
acattcatct gtgttgtgtt gtagcatgta tcagttacttc attcctttc acagcagaat 7380  
gatattccat tgtaaaacac tacattttt ttatccattc attagtttat aggccctttg 7440  
gctattgtga gtatgtttgc tgtggacatg tgcatacgag tatatttag aataaccttt 7500  
ttcagttatt tgggtatac acctaggatg agaattactg ggtcacatgg taattctgtt 7560

- 12 -

taattttctg aagaaccatc aaggtgatct ccacgggggc tgcaccattt ccaccagtaa 7620  
tgtaccaggg tcccaatttc tctacatcct tttcaatgct tggttatttc tgggttttt 7680  
ttttttcccc cccagtgtgg ccatcttact ggatgtgaag tggtatctca tggttttaat 7740  
ttgcatttac ctaatggcta attaacactg aggatctttt catgtgctga ttggctattt 7800  
gtatatgtca tttggagaaa tgtttattca agtcctttgt ccattttaa aattggcttg 7860  
tcttttgtt gagttgttagg gttcttata tattctggat attatthaat ttgtaaataa 7920  
ctcctcccat tctgtgggtt gtctttttt tgatagtgtc ctttgatgca caaaaatttt 7980  
agttttgctg aagtccaatt tatctttttt tcctttctt taggtgtcat atctaagaat 8040  
ccattgccaa acccaaggc atgaaggttt accgcatgtg ttttcttcta agagtttat 8100  
agttttcaact tatattnagg ctttgataaa ttttgagttt atttttgtat atgtgtgagg 8160  
caagtccaaac ttcattgttt tgtaactcaga tatccagttt tcccagcacc atttgttagg 8220  
ctgttttcc cctgttgaat ggtcttggta cctttgtaga aaatcaactg gccatagatg 8280  
tatggattta tttctagact ctcaattctt ttcatttttt tggtttgggtt gtttaagaaa 8340  
gggttgcatt ctttcgacag cccaggctgg agtacggtgg ctccatctt gctcaactgca 8400  
acctccgtct cctgggttca agcaattctc ccatctcagc ctcccaggta gctggacta 8460  
caggcgtgtg ctaccatgcc tggctaattt ttgtgtttct tggtagagat ggggtttcac 8520  
catgttggct aggctggtcc tgaattcgtg acctaagtg atttgctcac ctcggcctct 8580  
caaagtactg ggattacagg catgtgtgag ccactgcgcc cagccaatcc tattcatttg 8640  
atctatatgt caataccaca ctattnngt actgttactg tggcttactg tggttattgt 8700  
ggctttggag caaattttga aattccagat tgtgaggcct ccaactttgt tcttttttt 8760

- 13 -

ttttttagac gcagtcgc tttgtcgct atgctggagt gcaatggcgc gatctggct 8820  
cactgcaacc tccgccttct ggtttcaggt gattctcctg cctcagcctc ccgagtagct 8880  
gggattacag gcccggca ccacgcctag ctaattttc tatttttagt agagatgagg 8940  
tctcaccatg ttggtcaggt tggtctaaa ctccgtacct catgatctgc ctgcctctgc 9000  
ctccccaaagt gctgggatta cagggatgag ccaccgtgcc cagccaactt tgttttttt 9060  
taagatcggt ttggctgttt gaggtccctt gagattccat gtgaattata gcatcaactt 9120  
ccatttttg caaaaaaggc cattgggatt ttgacaggaa ttgcatttag taaattgctt 9180  
tggggagttt tgccatctta acaatattcg gtcttcaat ccatgaacat gggatgtctt 9240  
tccgtttatt tatgtcttta atttcttca gcaatgtttt gtagcttca atggacaaat 9300  
cttgcacctc ttggtaaat ctattccat gcatttatt ctttcgatg ttattataaa 9360  
tgaaattgtt tgaatttcct ttaagattt ttcattgctg gtatatacaa taatcagtt 9420  
tatagaaata caactgattt ttttgtttt atcttgtatc ctacaacttt gctgaatttg 9480  
tttcttagca tttttttctt tttttttttt tttttttttt ttttagacag agtctctc 9540  
tgttaccagg ctggagtgca gtggcatgat ctggctcac tgcaacctcc gcctccagg 9600  
ttcaagcgat ttttctgcct cagcctccca agtagctggg actgcaggtg catgccacca 9660  
tgcccagcta atttttgtat ttttagtaga gatggggttt cgccatgtt gccagtgtgg 9720  
tctcgatctc ttgacctcgat gatctgcccc cctcggcctc tcaaagtgtt ggtattacag 9780  
gcatgagcca ctgcgcctgg cctgtttctt agctttaata gttgtgtgtg tgtgtgtgt 9840  
tgtgtgtgtg tgtgtgtgtg tgtgtgtatt ctttaggatc ctctatataat aacatcatac 9900  
cgtctgtgaa gagaggttagc ttcccttcca atttggatgg cttttattha tttttcttgc 9960  
ctaattccctc tgattgaaac ttccagtaact atgttaaata gcagtagtgg agcaggcatc 10020

tttgtcttgt tcctgatctt agacagaggg cttcaatat tttaccattt agtataatgt 10080  
cagctgtggg gttaaatttt ttaacgcctt ttatcatgtt gagggagttc cttctgttc 10140  
ctagttgtt gagtgatttt atcacaaaag gctattgaat tttgtcaaag gcttttgtg 10200  
catcaactga gagatcggtt tttcccttc tctgctttt ctcccctct actggtagaa 10260  
aggacccacc taaagcaagc agtggcgcc cttagaggggt tacagccatg ctctccctg 10320  
agagcagttc ttggtttcaa cctgaggca gcgggtccgc ctgaggaaac caggtgtctg 10380  
gaaggtgaag gcttggag ctgagtagat gggcagtag gtcccagaga tatggccagc 10440  
cccagtcatg tcctgtctc tgtggagtcc cacagaggct gacgaggat gggggccctg 10500  
atagctggct acatgcaggc catgccctt ggcgggtggt ggcgtcagtc tggggcagac 10560  
ctcccatgct cacatagtgt gtcattcac ccagcaactgc cttaggttgg gtccttaga 10620  
atggtggctc ttaaacccca gcaagtatct gaaacactgg agggcttggt ccagcagatg 10680  
gctggggccc tcccagagtt tctgatccat gttgtcttgg gtagagactg ggaatctgca 10740  
tttctaatac attctcaagt gttgtggatg ctgctggctt gagaaccaca tccctagaag 10800  
cagagtctga gatggtgcag gcgatttcag atgaaccctg caagaggcac aggcagtgg 10860  
gagcgggcag agtgagcagc tgagcacaga tgtggatttg gaagtgtggc ctcagcctga 10920  
ttccatggag atctctgggg cgtaatgtc accacagggt tgccctgccc agaagcatgt 10980  
ggcctggctg ttacaggccc ttgtcagtca tggctctcct gggatgtgc aggtgaggtg 11040  
gcttcgtca ggagaaggc tctggtgac cagccagaaa aggggatcaa cggcatgcat 11100  
ggccagcacc tactgtgtgc cagggatggc ctcagcaactg tctgcacagc agtgagcaga 11160  
cgcggtgtgt ctcctggag ctggcatctt tttgagggag atagatgcta atcgggacag 11220

- 15 -

tctgttagcct cagggagaga agtgcttatct ggaaagatga agccaagggtg tgggctccag 11280  
ggggccccag gtgggagttat tttattttat ttttttggaa cagagtttca ctctgtcacc 11340  
caggctggag tgcagtggtg cgatcttggc tcactgcaac ctccacccct tgggttgaag 11400  
agattctcct ccctcgccctc ctgagtagct gggattacag gcacctgcca ccatgcccgg 11460  
ctaatttttgc tgtttttaat ggacaccaga ttccaccatg ttggccaggc tggcgtgaa 11520  
ctctggacct caagtaatcc gcctaccta gcctccaaa gttctggat tacagatgta 11580  
agccaccaag cctggctggg tgtgggatt ttagattaga tgaggaggac aggcctctct 11640  
gactggtttc cacctctaag tcctcatcca aagccttgtt ttatagatga gacagaggca 11700  
cagagaagtg aattctaaat tcacatagcc agtggcagaa cccagacttg gaccagttt 11760  
gggaacttct gagcctgtcc accccagttcc tagcctcacc cacagtgcc 11820  
cgagactatc agggagcctg acctgctgga tctggcagt cccaccgtgg catgctgcat 11880  
gtccccagaga aggtatctgt cagcagtgca gcacccccc 11940  
ctcggggct atccctggaa gtgttgtca gaaagtgaat ctccagatgt cacctgg 12000  
tgccctgagc tcctccttacc tgccacccctc tctgaccaca tagagcctgc tctagccc 12060  
gccctcttcc ctctcctccc ctcacccagg gacccggcac tagtccgccc caccactct 12120  
gtttatttct caccttggcc actgatgggt ggtttctctt agagcggtgc tgccctgtgg 12180  
aaccttctgc aatgatggaa atgctcagac ctgctctgtc cagtcacgtc gccactgg 12240  
gcatgtggct cttgaaatat ggagagtgta actgaggaac caaacttgaa tttttaaaat 12300  
tttgcgttgc ttacaatcac tcgtaagtag ccacccgtgg ctggcagcca ctggattgg 12360  
tggtgtgtgtt ctaggggttt ggcaaccaca tcactgcctt gtgcagaaac cactgctgca 12420  
ccaggagaag gccccaaatgc cagccctccctc ttcactgccc gaaggctgct gctccgctg 12480

- 16 -

ggggctcgtc tcgccaacgt tggcacagca aacacacata ctttctcctg tgggggctgg 12540  
tcctgctggc caagtcccggt gcatgctcct gggtggctgc acctggcccc tgcaccagg 12600  
caggtccaat ctgtggagga taccaaggaa cctcttttagt gttcccaagt gtgtcccatt 12660  
ccactgcagt tttgcagaag gtttagtgtgt gtgacttaaa aggcaaagag ggcaggcaga 12720  
tcttctgaca tctggggggg gcaaagtttag aatggaatat ttgctgcaga acttctcaga 12780  
gccttagca tgctaggatg tgctgcaaat ctccaggagg caggcggcat aagccatgct 12840  
tcccaaacga cttgccggtg gaagcctcct tgaggagtgc tgtgcgagac ccgtggctgt 12900  
ggagcacacg agagaatgcc tttctgtgg tttgtgtcca tgctggctc tcggctgcat 12960  
tgtcttccag tctgtgtccc ctgctggctt cccagggagg gagggaggct gtgactccat 13020  
tgctcccttc agcggctcggt ttgtttgtc attcggtcat ggaaaaccat ggttccatgc 13080  
cagccacacg cggggcctct gccgggcagt gggatgagtg tggtaacaa gaggagctga 13140  
tgacctcagg cagggacctt ccttcctctg ggtctgtccc gcaacataaca cacacgcaca 13200  
cacgcacacg gacataacctg tgcacacatg tatacacaag acacatacac acacatacat 13260  
acactcatgg gtgtgtcctg cagctgtctg gctgtgtgg tcccagctct tacactccca 13320  
ccccctccca gcccctgtga tgccctccatg ttaccgcccag agggcctggg cttgtggaag 13380  
tggtgccccg tgggcacctc tccttccca ccatgagtgg gaccctgctc actgccttct 13440  
ctaccagagt gagggagtga tgccagcttc ccccgcccttc agccgccttcc gccggcctgg 13500  
gctggtgccc atgggcattc cccagcagtg tgggcaggct gggtgccctgg cacccccagg 13560  
actatgacag aagcctcccc tggtgccag ggcctaagcc atgaggcccc tgctggggcc 13620  
tgacttaggg tggcgtcctgc cttttgtcccg gcctgagtg gcctggctac agcaccttt 13680

- 17 -

ggccctctga ggtcgtcac ccctctgcca tcacacccat ccctggccac cctctccctg 13740  
cctgctgcct gctgtctgtc attgaacatg ctcgtgttc tccccatccta aaactcctcc 13800  
tcctggttgg tgaacgcaat ggccacactt cccactttcc tctcatggaa tgtctgcagc 13860  
ttggtgccctc cctccacctg ctccttccag ccaccctctc tccacctggc ctcctgagca 13920  
ctgcacctta ggtctttcca catctcaccc tgtcccaggg aagcccttga tcgtccccag 13980  
gggtctctc tctgggcctt gcccttcagc atgggaagcc tgcagtccca acccagccct 14040  
tcacctccca ctctcccacc cctgttctga gtcaggatct cacttaaacc tcagtcgtct 14100  
cacctggctg ccccaggggc tgacttggcc catagagagc agaaccttagt gcccctctg 14160  
taccctgttt caggttcacc tccaagtgcc attaccctca cagggccctt acccgacacc 14220  
tggggccctct accccttgtc cctgcatgtc gcctgctaat acctgctctt cttaccaccc 14280  
cagacccttc ttatctcatg cttcctctctt agggctgcta cttctctatt cctgttcccc 14340  
taattggttc tccttgctgc agctagtgc gcttgggaca gcaccatcta tggttcccta 14400  
ctgcctgac gacaatgtgt gagcctgtgc taggagacca ggccctgtgt gataagctca 14460  
gcctgcccctg ttccagctgc acccacccttc tctagatcat ggactcactt ctctgcccac 14520  
agataacctt ttcccttgac ctctgcatct ggataactcc tattcactct tcacccctg 14580  
caaatgccat cacccccaga aagcctctct aataacccccc acccagttct cctcttcattc 14640  
accacactca tcacactgca aataagtgtc tgcaagtgtc ctggcatgag aatggggccct 14700  
ccagtgccca cctggggcac ctgcaggca cttagtaat atttacaaaag tgagtggtc 14760  
tgcctcgctt ggggtggggag cagggatgct ttttcagccca ggagatggct tggggtttgg 14820  
gttcagctgg gcagccagtg ccatggatat ttacctgggtc cacttggagg tcacagggca 14880  
cactctgtcc tgatcttagt gcagataacctt tcaggttacc gtagacccccc ccagcctcag 14940

cagctggaga tgagggcagt gcatccctt tgccaggaag gtccgattcc caatggacaa 15000  
agaggcaatg cagtgcgagg gagccagagg ccagggctcc cgtcccagct ctgtcagtga 15060  
ctcattgtgt ggccttggga agatcctcgc tgccctaggcc tcagtgtccc cttctgtaca 15120  
gtgggtggtc tagactaatt ttttatccca aagcagtccct agacctgcac tgctgacttg 15180  
gagccctctg cacctcctgt tctggcaca agagggcage caagggcctc agaacgctga 15240  
ggaaccctgg ccaacttagct ttaagaaaatg cattgtgtaa actgctctt actgagccca 15300  
gagcttgcca ggagcctggt agggttgtgg ctctggctct catttctacc aaaggaagtg 15360  
tgcttgacca gggagttcat ccaagggcac ctggaaactg tcctcaaggc atttcccggg 15420  
gaaccaattt ctcacggggtt gcctcagggtt ggggaagcgg aggccaacag cccctgtctt 15480  
tttccgcagt ggtcctttgc ttttgctacc atcacagaaa tcccccccgat tatcttcctc 15540  
cccaacttcc ttgtgcagag gaagggtctg agggcccttc ggacccagac aggaggaacc 15600  
ataatggtag gtgggggtggg ggggcatggc tgggtgggg gccccacac cccagggtcc 15660  
ttctcacctc cttgccctg gaatgcctc ctcccactta gtagttgaac agaatcctaa 15720  
atattcctca aggctcgca acaatgaccc ttttccaaa agcctttttt ccccatctt 15780  
ggacatcaga attctttctt catcgttctt tctcctatga cctcctattt gttaccgtaa 15840  
ttgctagtat ataataacc tctccaccca ccaaagcggc tttccatgc cttatggcttt 15900  
aaggcacacc ccctcaccag tttttttttt ctttctttctt tttttttttt tgagttagat 15960  
ctcggtctgt cgcccaggct ggagtgcagt ggtgtatct tggctcactg caacctctgc 16020  
ctccttaggtt caagcgattc tcttgccatca ggctcctgag tagctggac tacaggttt 16080  
cgccaccatg cctggctaat ttttgtatcc ttagtagaga cggggtttta ccatgttggc 16140

- 19 -

caggctggtc ttgaactcct gacctaaat gatccactca cttggcctc ccaaagtact 16200  
gggattacag gcttgagcc accatgccca gcctaatgc accaaaaatt aagatggaga 16260  
actgatcctc catgacttca gtatgtata agcctccacg tctccccac tgccgggtgt 16320  
gcaacaaaga atccccacag caaaatttagg tttcacattt tttgtgttgt ttttttaaaa 16380  
aatgtgcca cacactgcct agttatttgg agatagagga atgttcaca tgcaaatgt 16440  
tgaggatcta acccagccct ggatcaactac ctactgatcc cctacagttc tgttatgtt 16500  
gtaaatttgt acttttcct ttagcttagt agaatattac tgcccatccc caaaactatg 16560  
atttcctgga agatttcagt atttagtcta ctatatttct tttttgttt tttttttttt 16620  
tttttttag acagagtctc actctgtcct ccaggctgga gtgcagggt gtgacccgtt 16680  
ctcaactgcaa cctctgcctc ctgggttcaa gtgattctcc tgcctcagcc tcccgagtag 16740  
ctgggattac aggcacacgc cactctgcct ggctaatttt tgtatttta gtagagacgg 16800  
ggttcacca tggcggcag gctggcttg aactcctgac ctcaagtgtat ccgcctgcct 16860  
cgccctccca aagtgtggg attacaggcg tgagccactg cgccctggccc agtctactgt 16920  
atttctgtga gcaaaacttt gcctatttc ctttgaaag ccataatcaa attattgtca 16980  
gctcatatgt gatggatgat aagtactttt atttttcca gtttccttgc acaatttcaa 17040  
aggtgcttat gcactgtaca tctcatatgc cagccaagct ggcacttact tcctggactg 17100  
ttgcttgggg tagggagttc cttctataacc cctgccttgt agctcagctc atccttcccc 17160  
cagagctggc tagaaggcagt gtttatggaa tgagtgcattg aatcagtgaa tgaatgactg 17220  
gtggatcgcc tgcctgcgcc ccctcacccct ctgcttgcct ccaaaggcgg ggaagctggc 17280  
tgtggagcga ggctggccca tcaacgtggg tgagtgcctt gatatgcctc ggaaatgtcc 17340  
agcccccgtt ggtggactg gcctgaaagg gggctggggg agggcggag gatcctggag 17400

tgcccgactg tgaattcaga agctctggtt ttcccaagtc acccttagcct ccttgcggag 17460  
tggcctggag gttgtatgtgt agcctcctag gtacctggga gagactgacc agtgcctcca 17520  
tctgacgtgg gatccttgtc taaggaggc cccgggtggt tccccagccc cctctttgcg 17580  
tacttccggc ggcaggggagc ttccctccctt ccagagagcg tgtgccatcc ttggcagct 17640  
cagcatggtc tgaagcctgc cttgtgtt ccctgaagga ctccacctgt gtcctggggc 17700  
ccaggacagc ccacagaggc ttggtcatgt tgggttgggt gggcacatcc tgggtcaata 17760  
ccaccacctt ctcaagggtc cagagggccc gtgctccca gcccccttga atctcccaca 17820  
agattggctc atgggagggc tgacacggag tctcccttgt ccctgtcatt gtccttcctg 17880  
gaggcacagc acttgacaat ttacaaagct cttttcacc aggcttttt tttcttttc 17940  
gagacgtagt ttcaacttttgc ttgcccaggc tggagtgcaa tggcgcgatc tcggctcacc 18000  
gcaacctccg cctcccaggt tcaaacaatt ctccgcctc agcctcctga gtagctgaga 18060  
ttacaggcat gcaccaccat gcccggctaa tttgtatTT ttagtagaga cagggtttct 18120  
ccatgttgggt caggctgggt cttgaactcc cgacctcagg tgatacgccc acctcgctcg 18180  
gcctccaaa gtgctaagat tacagacatg agccaccacg cccggccttc acccagactc 18240  
ttatTTgagc tgggcataat tgtaGGGCT gtctcaCTGA tgaggAAATg gCCATGGAAA 18300  
gatgcgtact ggatcggtgt a gacccctaaa gcagggtccc ccagcTTTg gctctgaact 18360  
ctgcaggggaa gagtccacct tggGCCACTG cacagttgag gggagccccca ctctgcagg 18420  
gctgggtctc ttccatTTGC gtattaccag gtgcctAGCA ttcAGTCTGG catAGTAATG 18480  
atgttatggt actctgctgc acaaACCCGG gagtgatCTG tgccCTGCGT gtctacAGCA 18540  
gggttccgag gagggcctgg atggccCTCC ccatggcagg tgTTactGCC tggtAGAGGT 18600

taagagcctg gatcctgatc caccctgggt ttgatcctgg ttctgcatt acctggctgt 18660  
gtgaccctgg gcaagttgct gacccctct gtgggtcagt ctccatct gtaaaatggg 18720  
gatggtgatg ctaatcccc tcctcggtt ggagggagtc ttcagcaagc tcagttgctc 18780  
agtcaggtgt tcactgtggc tgtcttctca tcattaggag ccaacagtag cctcctgggg 18840  
ggtgggagag gcaagttcct ggtatccatg gggccagctg cacactgtct gacggagcag 18900  
ttgtgggct caatttcaga gggcctctgc aattcaggcc atcccagggg ctgcaggga 18960  
gggggtatct atggcccta gggctctgag gctgtgtctc agggttgagg ggtgatggat 19020  
cccgccctct agggccctcc tcgtggctgt aggcagtcat gaccagcaga gggtgccctt 19080  
cctgaccacc cgctttggcc actggcagaa tccgtgtggc cccatacca ccactccctc 19140  
ctggagtggg gagccacatg gagccaggcc cagcttggtg gggacaagga gcagctttct 19200  
gcttctggaa tgatgagcta tctgttgctt aggggtgtga gtggcaactga ggacttgctg 19260  
gggacaccct gaagatgtgg ctgccttctg gcctggggat ggtgacatgc cccagcactc 19320  
agcttagttt gccaacccag agtccgaggc acaggttcct gagagctgag cagggaggat 19380  
gctggggag gtgaaggat ggaggagctc ctggactgag cctggagcc tggctctgag 19440  
cagcaccgct ctctgcctt ccgcaggggg tggcttccac cactgctcca gcgaccgtgg 19500  
cggggcttc tgtgcctatg cggacatcac gctcgccatc aaggtgtgtc tatgagcaag 19560  
tgggtctcg cctccaagag ccctcctgga atccctccca tagctccaaa ttaactgttc 19620  
tcaccctgaa ttatagacaa gggccatag ctggagcagg gagggggctt gtttgggttg 19680  
ctcagccagg ctggaactga atccagatct gacacttgct cctttccat gttgcttaga 19740  
agggttgct gtggtggaaag ggagttattc cagectccca cagagccagg ggactagaga 19800  
gggtcaggat ctgctgtata gccacatatt aagttgtagg aagaaggca tggctggcaa 19860

- 22 -

agggagtagg gagtgaaag aatgatggtg ctgatagcac ctggcagttc tgcatactcc 19920  
aaccccgct gtgtccagg acttactccc tgaatcctcg cagacagaca ggggcccaca 19980  
gaggtgaggg catgaaata gcagggcag aattggcgct ggccctgtt ctgtgggccc 20040  
ccacaactcc cctgccactc tgtgcctggc cttgtgtgg gcatcaggaa ctgactgacc 20100  
tgttcctatg tgtgcctgtc ctcatggggc acatagactg atgggggaa gcaggccatt 20160  
aggagaaggg ggaagcacag gagacccccc tggggaggag ggaatgaagg cttcctggaa 20220  
gagggggcat ttaggacttg gcctttagg ataaggcaga ggttggggac tgaagtccca 20280  
gggctgtggg gattctctcc ttaacccta cacatccct agggaatctg gaaaaatcca 20340  
gggcctgagt gacccactta ctcctgacc tatgaccctt cagggcacag gacatgcccc 20400  
ctcctccagg gacccccc tgaccaccc tcgtatgcac acatggagcc ccacagctgg 20460  
agctgcacag ctctccctgg caagtgacat cttgtgtgg tggcctgatt acccacaagc 20520  
attaggcccc ctcctccgc ctcgcggc cagctgggag ttgctgttagg gctgggtcct 20580  
ctgtccggcc cagatcctca tgtctaccct ctcctccctg gcagttctg tttgagcgtg 20640  
tggagggcat ctccagggtc accatcattt atcttgcattc ccatcagggtg agtgcctgc 20700  
aggggctgga ctcttagggg acctgcacc cccagttcca gaatctccc gggcaggag 20760  
agtctccctc ctcatgtccc cacggcttc acggcttctg tcttctgtct ctggggctac 20820  
aaatgcaggc tctgtctttt tcactctgtc caggacagcg ggtcctccctc attgtcccg 20880  
agggtcctcc ctccctccctc ctgactgccc ccacatgagg ctcttcctga agccactct 20940  
gatgggactg ctctcggttg cagagctctg ctgtgggtcc ccattgctta tgaataattt 21000  
ggggcactgc cccctgccc gagctgctga gcactggcca ctcgtccctc aggccggatgc 21060

ccacacacat ggcttggctc gggcacctgg ggtcaccatt taagaactcg gcgcctaggg 21120  
agtaaaagtgt caaagcagag ggttacacctcc tcctcaggac ccctaattgag gccagtgcct 21180  
ctggtcagac agggagggga cccagtgggc tccggaaaggc acccccctgc accattactg 21240  
ctgtggcttt gtgcttagttg gggccctgcc ttgggttctt gcgaccccgaa actcctgagc 21300  
caggtcacat gtggacagtc cttaacagtt tgctttcac atccctgatc ccaaccagtc 21360  
ccaccacaga cttgagaggg tggcagagcg ggatttcttc ctctgatagg gaaacctaaaga 21420  
gcactgggct tgctcaagcc catgctagaa ggtgtcgaaaa cctggttta aggttgaatc 21480  
ccagctctgc cccttaaacag tcatgagacc tgctgcccc gagagcaggg cgtgctgccc 21540  
tggcaaattgg ggagtttcct gaggggtggg tgggtggcag agccccagcc ttgcctaggg 21600  
cacctacccg agagcggcta ctgtgacctc cccacagggc aatgggcatg agcgagactt 21660  
catggacgac aagcgtgtgt acatcatgga tgtctacaac cgccacatct acccagggga 21720  
ccgctttgcc aagcgttaagc tgctgcccc accctcatct tgggtgtgtc cttgtggatg 21780  
aggctctctc ctgagtgtct cctgtctgct aggccctgca gaagccactg cagtggttca 21840  
tagcatccct gtgaggtgat cctttccatt ttacagatga ggaaaccgag acctggagaa 21900  
gtcactcgac ccacccaaga tcacataacc cttacaataa acatgcattt gtctggcaaa 21960  
aaacagggaaa gaatgaaaga aaaaaaaagaa aaataggata aatttggaaaa tacgaaataa 22020  
gaaataaaatt cacataggct gggcgccgtg gtcacgcct gtaatcccg cactttggga 22080  
ggctgaagtg ggcggatcac ctgaggtcg gatggatgag ccagcctgac caacatggag 22140  
aaaccccatc tctactaaaa atacaaaatt agctggatgt ggtggcgcat gcctgtaaatc 22200  
ccagctactc gggaggctga gacaggagaa ttgcttgaac ctgggaggcg gaggttcgg 22260  
taagccgaga tcgcgcatt gcaactccagc ttggcaaca agagcgaaac tccatctcga 22320

- 24 -

aagaaagaaa gaaattcatg tataatcggtt aaaatgaaaaa tgcattaaac tcatcaatca 22380  
aaaggcagag actctcagat gagattaaa aacagggctg ccaccttgc agtagggga 22440  
cattttgca ccagtcacga tgagtctggt gtggataagt cagcagctag tatggccaa 22500  
ggaaccaatt tctgaacaga acctcacatg tgctgagcct gggcttaagg gcagggcagg 22560  
gtgtccatgt gtgtaggcaa gacccagagg aggcatgtaa atctgacatt gccgacacag 22620  
atctccacac ccccagggca gtgtctcage ttcaagtgcctt cttctctcct ttgagtcccc 22680  
cttttgca cttttgcag ctcttggtgc tcttttacc ttagtttgg gtggaatgag gctgagcagt 22740  
gctgaatctg acagaccagt ttccagtctt gcctggtgac cacagtcttgc tcctgagcct 22800  
cagttccct tctctataaa ttgaggccat ccatgtctctt ctccagagg ccatcaggcg 22860  
gaaggtggag ctggagtggg gcacagagga tgatgagttac ctggataagg tggagaggaa 22920  
catcaagaaa tccctccagg agcacctgcc cgacgtggtg gtataacaatg caggcaccga 22980  
catcctcgag ggggaccgccc ttggggggct gtccatcagc ccagcggtac gtcctgaccc 23040  
ttggggccac gggaggggtct gctctatggc ctcagcagca gcagggaaagg tggggccct 23100  
catgtcaggg aggagatggc ctgaagcaac agcagtttgg agcaggctt gcccgtcagc 23160  
aggacttcctt gacaccatgg gggtctggcc tgcctgagtc accctccctt tccccctaaca 23220  
gggcacatcgta aagcgggatg agctgggtt ccggatggtc cgtggccgccc gggtgcccat 23280  
ccttatggtg acctcagggcg ggtaccagaa gcgcacagcc cgcatcatttgc tgactccat 23340  
acttaatctg ttggccctgg ggctcattgg gcctgagtc cccagcgctt ccgcacagaa 23400  
ctcagacaca ccgcgtcttc cccctgcagt gccc 23434

**THIS PAGE BLANK (USPTO)**

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 February 2003 (20.02.2003)

PCT

(10) International Publication Number  
**WO 03/014340 A3**

(51) International Patent Classification<sup>2</sup>: C12N 9/16, 15/00, C12Q 1/68, G01N 33/50, C07K 16/40

(74) Agent: GROS, Florent; Novartis AG, Corporate Intellectual Property, Patent & Trademark Department, CH-4002 Basel (CH).

(21) International Application Number: PCT/EP02/08654

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LT, LU, LV, MA, MD, MK, MN, MX, NO, NZ, OM, PH, PL, PT, RO, RU, SE, SG, SI, SK, TJ, TM, TN, TR, TT, UA, US, UZ, VN, YU, ZA, ZW.

(22) International Filing Date: 2 August 2002 (02.08.2002)

(84) Designated States (*regional*): Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR).

(25) Filing Language: English

(26) Publication Language: English

Published:  
— with international search report

(30) Priority Data:  
60/309,957 3 August 2001 (03.08.2001) US

(88) Date of publication of the international search report:  
27 November 2003

(71) Applicant (*for all designated States except AT, US*): NOVARTIS AG [CH/CH]; Lichtstrasse 35, CH-4056 Basel (CH).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(71) Applicant (*for AT only*): NOVARTIS PHARMA GMBH [AT/AT]; Brunner Strasse 59, A-1230 Vienna (AT).

(72) Inventors; and

(75) Inventors/Applicants (*for US only*): ATADJA, Peter, Wisdom [CA/US]; 18 Eastbrook Road, Parsippany, NJ 07054 (US). CUETO, Maria [US/US]; 99 Clifton Terrace, Weehawken, NJ 07086 (US). GAO, Lin [US/US]; 8 Millstone Drive, Morris Plains, NJ 07950 (US).



A3

WO 03/014340

(54) Title: HUMAN HISTONE DEACETYLASE-RELATED GENE AND PROTEIN HDAC10

(57) Abstract: Disclosed is an HDAC related gene and gene product. In particular, the invention relates to a protein that is highly homologous to known HDACs and referred to herein as HDAC10, nucleic acid molecules that encode such a protein, antibodies that recognize the protein, and methods for diagnosing conditions related to abnormal HDAC10 activity or gene expression.

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP 02/08654

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 C12N9/16 C12N15/00 C12Q1/68 G01N33/50 C07K16/40

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 C12N C12Q G01N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EMBL, EPO-Internal, WPI Data, PAJ, MEDLINE, EMBASE, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL 'Online! 29 September 2000 (2000-09-29) WATANABE ET AL.: "Homo sapiens cDNA: FLJ22237 fis, clone HRC02058." retrieved from EBI Database accession no. AK025890 XP002225132 Polypeptide and cDNA with 100% identity over full lenght. abstract</p> <p style="text-align: center;">— —/—</p>	1-22

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

## \* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*& document member of the same patent family

Date of the actual completion of the international search

Date of mailing of the international search report

7 January 2003

05/02/2003

Name and mailing address of the ISA

European Patent Office, P.O. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Friedrich, C

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP 02/08654

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE EMBL 'Online! 29 March 2000 (2000-03-29) WANG ET AL.: "Homo sapiens chromosome 3 clone RP11-449E21 map 3p." retrieved from EBI Database accession no. AC027124 XP002225133 genomic sequence abstract</p> <p>-----</p>	1-22
P,X	<p>DATABASE EMBL 'Online! 31 January 2002 (2002-01-31) MEYERS ET AL.: "Human histone deacetylase 47508." Database accession no. AAM51008 XP002225134 abstract</p> <p>-----</p>	1-22
P,X	<p>-&amp; WO 02 08273 A (MEYERS RACHEL A; MILLENNIUM PHARM INC (US)) 31 January 2001 (2001-01-31) examples 1-5</p> <p>-----</p>	1-22
A	<p>FISCHLE WOLFGANG ET AL: "A new family of human histone deacetylases related to Saccharomyces cerevisiae HDA1p" JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BALTIMORE, MD, US, vol. 274, no. 17, 23 April 1999 (1999-04-23), pages 11713-11720, XP002159645 ISSN: 0021-9258 the whole document</p> <p>-----</p>	1-22
A	<p>GROZINGER CHRISTINA M ET AL: "Three proteins define a class of human histone deacetylases related to yeast Hda1p" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE, WASHINGTON, US, vol. 96, no. 9, 27 April 1999 (1999-04-27), pages 4868-4873, XP002159644 ISSN: 0027-8424 the whole document</p> <p>-----</p> <p>-/-</p>	1-22

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/EP 02/08654

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	KAO HUNG-YING ET AL: "Isolation and characterization of mammalian HDAC10, a novel histone deacetylase." JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 277, no. 1, 4 January 2002 (2002-01-04), pages 187-193, XP002225131 January 4, 2002 ISSN: 0021-9258 figure 1	1-22

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.1

Although claims 11-14 are directed to a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

---

Continuation of Box I.1

Rule 39.1(iv) PCT - Diagnostic method practised on the human or animal body

---

Continuation of Box I.2

Present claims 10, 11, 13, and 22 relate to a compound defined by reference to the term HDAC10. The use of this parameter in the present context is considered to lead to a lack of clarity within the meaning of Article 6 PCT. It is impossible to compare the parameter the applicant has chosen to employ with what is set out in the prior art. The lack of clarity is such as to render a meaningful complete search impossible. Consequently, the search has been restricted to HDAC10 according to SEQ ID No.1-4.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/EP 02/08654

### Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: because they relate to subject matter not required to be searched by this Authority, namely:  
see FURTHER INFORMATION sheet PCT/ISA/210
2.  Claims Nos.: because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:  
see FURTHER INFORMATION sheet PCT/ISA/210
3.  Claims Nos.: because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

### Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1.  As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2.  As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3.  As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4.  No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

#### Remark on Protest

The additional search fees were accompanied by the applicant's protest.  
 No protest accompanied the payment of additional search fees.

**INTERNATIONAL SEARCH REPORT**

Information on patent family members

International Application No

PCT/EP 02/08654

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 0208273	A 31-01-2002	AU 7710501 A WO 0208273 A2 US 2002164752 A1	05-02-2002 31-01-2002 07-11-2002

**THIS PAGE BLANK (USPTO)**